The Journal of Experimental Education

A periodical report of scientific investigations relating to child development, curriculum, learning, teaching, supervision, measurements, statistics, and experimental techniques.

Volume XXIII | June, 1955 | Number 4

CONTENTS

Page

Prediction of the Probability of Graduation of Engineering Transfer Students — *J. Stanley Ahmann* 281

Characteristics of the Effective Teacher Based on Pupil Evaluations — *Percival M. Symonds* 289

Significant Characteristics of College Students with Varying Flexibility in Reading Rate:
I. Eye-Movements in Reading Prose
II. Motor and Perceptual Skill in "Reading" Material Whose Meaning is Unimportant — *Frank Laycock* 311

An Evaluation Instrument for the Analysis of Teacher Effectiveness — *Granville B. Johnson, Jr.* 331

Attitudes and Professional Relationships of Teachers: A Study of Teacher Morale — *E. C. Hunter* 345

The Influence of a Child Purpose Upon the Perseverance of Young Children — *Louis Ada Wilson* 353

What is an I.Q.? — *Florence Brumbaugh* 359

Partial Reinforcement Related to "Free" Responding in Extinction with Pre-School Children — *Nicholas A. Fattu, Edmund V. Mech, Donavon Auble* 365

An Appraisal of Teacher–Counselor Adviser Teaching — *Russell L. Jenkins* 369

Peer Group Rating of High School Pupils with High and Low School Marks — *Evan R. Keislar* 375

Regression Coefficients vs. Correlation Coefficients — *John Schmid, Jr.* 379

$5.00 A YEAR | PUBLISHED QUARTERLY | $1.50 A COPY

Published by Dembar Publications, Inc.,
Madison 3, Wisconsin.
Entered as second-class matter October 17, 1938 at the post office at Madison, Wisconsin, under the act of March 3, 1879.

EDITORIAL BOARD

A. S. Barr, Chairman, Professor of Education, University of Wisconsin, Madison 6, Wis.

Arthur T. Jersild, Professor of Education, Teachers College, Columbia University, New York City. Editorially responsible for materials on child welfare, guidance, and development, published each December.

Palmer O. Johnson, Professor of Education, University of Minnesota, Minneapolis, Minnesota. Editorially responsible for materials on measurements, statistics, and methods of experimental research, published each March.

H. H. Remmers, Professor of Educational Psychology, Director Division Educational Reference, Purdue University, Lafayette, Indiana. Editorially responsible for materials on learning, teaching and supervision, published each September.

J. Wayne Wrightstone, Director, Bureau of Educational Research, Board of Education of the City of New York, Brooklyn, New York, 110 Livingston Street, Brooklyn, New York. Editorially responsible for materials on curriculum construction, published each June.

CONTRIBUTING EDITORS

Emmett A. Betts, Director, Betts Reading Clinic, Haverford, Pennsylvania.

Leo J. Brueckner, Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Oscar K. Buros, Associate Professor of Education, Rutgers University, New Brunswick, New Jersey.

Guy T. Buswell, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Harold D. Carter, Associate Professor of Education, University of California, Berkeley 4, California.

Leslie L. Chisholm, Associate Professor of Education, State College of Washington, Pullman, Washington.

Herbert S. Conrad, Technical Consultant, College Entrance Examination Board, Princeton, New Jersey.

Stephen M. Corey, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Robert A. Davis, Professor of Education, Director of Bureau of Educational Research, University of Colorado, Boulder, Colorado.

Hari R. Douglass, Director of College of Education, University of Colorado, Boulder, Colorado.

Harold A. Edgerton, Director, Occupational Opportunities Service, Professor of Psychology, Ohio State University, Columbus 10, Ohio.

John C. Flanagan, Professor of Psychology, University of Pittsburgh, Pennsylvania.

Carter V. Good, Dean, Teachers College, University of Cincinnati, Cincinnati 21, Ohio.

Robert W. B. Jackson, Assistant Professor of Educational Research; Assistant Director, Department of Educational Research, Ontario College of Education, University of Toronto, Toronto, Canada.

Harold E. Jones, Professor of Psychology and Director, Institute of Child Welfare, University of California, Berkeley 4, California.

Noel Keys, Professor of Education and Lecturer in Human Relations, University of California, Berkeley, California.

D. Welty Lefever, Professor of Education, University of Southern California, Los Angeles, California.

Edward A. Lincoln, Consulting Psychologist, Halifax, Massachusetts.

Irving Lorge, Professor of Education, Executive Officer, Institute of Psychological Research, Teachers College, Columbia University, New York 27, New York.

A. R. Mead, Director of Educational Research, University of Florida, 330 P. K. Younge Building, Gainesville, Florida.

T. E. Newland, Lt. Comdr., USNR, 2702 Wisconsin Avenue, N. W., Washington 7, D. C.

C. W. Odell, Professor of Education, University of Illinois, Urbana, Illinois.

Willard C. Olson, Professor of Education, Director of Research in Child Development, University of Michigan, Ann Arbor, Michigan.

Valworth R. Plumb, Chairman, Division of Education and Psychology, University of Minnesota (Branch), Duluth, Minnesota.

S. L. Pressey, Professor of Educational Psychology, Ohio State University, Columbus, Ohio.

Clarence E. Ragsdale, Professor of Education, University of Wisconsin, Madison, Wisconsin.

William Reitz, Associate Professor of Education, College of Education Examiner, Wayne University, Detroit 2, Michigan.

Henry D. Rinsland, Professor of Education and Director of Educational Research, The University of Oklahoma, Norman, Oklahoma.

Robert T. Rock, Jr., Professor of Psychology, Head of Dept. of Psychology, Graduate School, Fordham University, New York City.

Philip J. Rulon, Professor of Education, Harvard Graduate School of Education, Cambridge 38, Massachusetts.

Douglas E. Scates, Professor of Education, Duke University, Durham, North Carolina.

John Schmid, Board of Examiners, Michigan State College, East Lansing, Michigan.

Harold Seashore, Director, Test Division, The Psychological Corporation, New York 18, New York.

David Segel, Educational Consultant, Specialist in Tests and Measurements, Federal Security Agency, U. S. Office of Education, Washington, D. C.

Paul W. Terry, Professor of Educational Psychology, University of Alabama, University, Alabama.

Helen Thompson, Associate Attending Psychologist, New York Post-Graduate Hospital, 303 East 20th Street, New York 3, N. Y.

Robert L. Thorndike, Associate Professor of Education, Teachers College, Columbia University, New York City.

Herbert A. Toops, Professor of Psychology, Ohio State University, Columbus, Ohio.

Maurice E. Troyer, Director, Bureau of School Services, Syracuse University, Syracuse 10, New York.

Helen M. Walker, Professor of Education, Teachers College, Columbia University, New York City.

Beth L. Wellman, Professor of Psychology, Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

Guy M. Wilson, Emeritus Professor of Education, Boston University, 33 Pine Street, Wellesley Hills, Massachusetts.

Paul A. Witty, Professor of Education, Director of Psycho-Educational Clinic, School of Education, Northwestern University, Evanston, Illinois.

Ernest R. Wood, Professor of Education, New York University, New York City.

D. A. Worcester, Chairman Department of Educational Psychology and Measurement, University of Nebraska, Lincoln, Nebraska.

DEMOCRAT PRINTING COMPANY
MADISON, WISCONSIN

Journal of Experimental Education

Volume XXIII June, 1955 Number 4

PREDICTION OF THE PROBABILITY OF GRADUATION OF ENGINEERING TRANSFER STUDENTS

J. STANLEY AHMANN
Cornell University

THERE IS little doubt in the minds of many that the educator of today is deeply interested in the prediction of the future. Even a casual examination of the literature readily demonstrates the willingness of investigators in education to devote considerable time and effort in attempts to determine their ability to predict the educational scene of tomorrow from known data of today. Specifically, repeated studies have been made to determine estimates of the future academic achievement of students, those estimates being based upon such present measures as scholastic aptitude, mechanical aptitude, interest, and previously demonstrated scholastic ability.

It is evident that the results of these efforts to predict, most of which have been based upon principles of correlation and linear regression, have assisted the instructors and counselors of students. This is true for the most part in engineering education as well as in other areas. Needless to say, however, the degree of assistance has varied greatly. For example, predictive studies concerning certain groups of students are comparatively few in number. One such group is the transfer student group, especially the transfer students in engineering.

Although education research and editorial comment reported in the higher education journals tend to exclude the question of the position of transfer students in colleges and universities, the importance of this group of students is steadily increasing in the opinions of many. In engineering, for example, instructors and administrators alike are becoming more and more aware of the impact on engineering colleges made by the undergraduate students who first began their college studies at institutions of higher learning other than the one from which they ultimately hoped to graduate. The academic and social adjustments to be made by such students who transfer from one institution to another are often genuine problems for the instructors and counselors of both institutions concerned.

The Division of Engineering at the Iowa State College enrolls sizeable numbers of transfer students each academic year. Estimates have been made that as many as 40 percent of the entering engineering students at the beginning of a fall quarter had received college credit from other institutions of higher learning. The percent of the engineering graduating class who were originally transfer students has been as large as 50 percent on occasion.

In the light of these figures alone, further investigating of the collegiate academic careers of the transfer students entering engineering would seem warranted. In addition, college officials and transfer students have repeatedly inquired about the likelihood of a student's academic success in engineering after transfer. The answers to such questions have been shrewd guesses at best since few adequate tools have been developed which will reduce to administrative routine the prediction of a transfer student's academic success in the Engineering Division.

It was the purpose of this study (1), therefore, to devise an instrument to predict the probability of academic success in engineering of students who transferred from other institutions of higher learning. For the purpose of this study, academic success has been defined in terms of whether the transfer student eventually graduated in engineering at the Iowa State College. It should be noted that this dichotomous criterion of graduation-nongraduation differed from the usual numerical criterion of course marks or cumulative grade point averages used in regression studies (4). However, the definition here chosen is considered to be reasonable in light of the fact that some engineering faculties seem to be more interested in whether a student graduates in engineering than whether he attains some numerical academic average above the minimum required for graduation. Furthermore, in view of the present day shortage of trained engineers, certain industrial concerns are inclined to consider a student's ability to graduate in engineering from a specific institution as adequate demonstration of academic achievement. Lastly, such a simple definition of academic achieve-

ment appeals to the entering transfer student who some day may be counseled on the basis of these results.

Prediction of the probability of graduating in engineering was based upon various combinations of test scores compiled by the psychological testing program, the high school grade point average, a rating of the student's previous collegiate academic achievement, and the first quarter grade point average in engineering at the Iowa State College. In all, six prediction variables were considered. Four of the variables were assembled as a matter of routine and were available prior to the date on which the transfer students began their studies at the Iowa State College. They were the linguistic raw scores and quantitative raw scores on the American Council on Education Psychological Examination 1945 Edition, the high school grade point averages, and the raw scores on the United States Armed Forces Institute Test on Correctness and Effectiveness of Expression, College Level.

The fifth prediction variable which also was available prior to the beginning of the transfer student's Iowa State College studies, was a rating of each student's prior college achievement as represented by the transcript of his credits. This rating was made on a six point scale by clerks in the admissions branch of the registrar's office, and was based upon the type of courses previously completed by the student, the final marks he received, and the engineering curriculum he was entering at the Iowa State College.

The sixth prediction variable was the first quarter grade point average at the Iowa State College. Because the transfer students entered the Engineering Division with a variety of collegiate backgrounds, the first quarter grade point average by no means represented academic achievement in a uniform set of courses. The five point letter marking system used in both the high school marks and the first quarter college marks was converted to a 4, 3, 2, 1, 0 scale.

Although the prediction variables resemble closely those used in regression schemes, the criterion, the tendency to graduate in engineering, was a dichotomy rather than the usual numerical variable. Consequently the regression technique was not applicable, and the discriminant function as reported by Fisher (2) was considered. Fisher developed the discriminant function for the purpose of finding the weights to be assigned to a series of variables that would produce maximum separation between two groups, each of which is assumed to be normally distributed. In the case of the present criterion maximum separation was also desired but within a single normally distributed group which had been divided into a dichotomy.

The foregoing assumption that the tendency to graduate in engineering is a single normally distributed variable necessitated a modification of Fisher's discriminant function, such a modification having been reported in a forthcoming textbook by Wert, Neidt, and Ahmann (6). This modification yielded a prediction equation resembling greatly the regression approach, with the major exception that the prediction is made in terms of chances in 100 of graduating in engineering rather than in terms of achievement scores. The general form for the modified discriminant function of m prediction variables is:

$$V_O = a_1X_1 + a_2X_2 + a_3X_3 + \ldots + a_mX_m + C$$

where

V_O = sigma units
a = coefficients
X = prediction variable scores
C = constant

Substitution of the appropriate values of the prediction variables yields a prediction in terms of sigma units. By consulting a table of the normal curve, the sigma units are converted to chances in 100.

To determine the size of the coefficients which would yield the best possible prediction equation, data was gathered for 804 male engineering transfer students who enrolled since September, 1946, and who at a later date, permanently left the Engineering Division either by means of graduation or dropping out. Of the 804, over three-fourths were veterans of World War II. In general, most of the students had attended only one college prior to enrolling at the Iowa State College, and the college was usually located in Iowa or an adjacent state and enrolled less than 2500 students. Approximately one-third of the students first attended junior colleges before transferring whereas the remainder first matriculated at colleges other than junior colleges.

Since earlier research (3) has indicated that veteran students tend to surpass non-veteran students in academic achievement, the original group of 804 was so divided. In addition preliminary investigation of the available data revealed the possibility that, when individual differences are ignored, those transfer students who first attended junior colleges experienced greater difficulty in graduating in engineering than those who first attended colleges other than junior colleges. Although the reasons for this situation were not determined, the veteran and non-veteran transfer students were further stratified on the basis of college status. The resulting four subgroups were:

1. Veteran students from junior colleges ($k = 207$)

TABLE I

TENDENCY TO GRADUATE IN ENGINEERING OF 804 TRANSFER STUDENTS DIVIDED INTO FOUR SUBGROUPS

Subgroup	Graduates		Nongraduates		Total	
	No.	%	No.	%	No.	%
Veteran, Junior College	52	25.1	155	74.9	207	100.0
Non-Veteran, Junior College	8	13.1	53	86.9	61	100.0
Veteran, Other College	159	38.4	255	61.6	414	100.0
Non-Veteran, Other College	27	22.1	195	77.9	122	100.0
Total	246	30.6	558	69.4	804	100.0

TABLE II

WITHIN SUBGROUPS BISERIAL CORRELATIONS BETWEEN THE CRITERION AND ALL PREDICTION VARIABLES

Prediction Variable	Biserial Correlation
Linguistic Score on ACE Test	.187
Quantitative Score on ACE Test	.296
High School Average	.312
USAFI English Score	.281
Admissions Office Rating	.396
First Quarter Average	.655

TABLE III

MULTIPLE BISERIAL COEFFICIENTS OF CORRELATION RESULTING FROM VARIOUS DISCRIMINANT FUNCTIONS

Combinations of Prediction Variables	Multiple Biserial Coefficients
Q-score, High School Average	.380
Q-score, High School Average, USAFI Eng.	.399
Q-score, High School Average, USAFI Eng., Rating	.480
Q-score, High School Average, USAFI Eng., Rating First Quarter Average	.686

TABLE IV

CHANCES IN 100 OF GRADUATING IN ENGINEERING FROM PRE-MATRICULATION DATA

H.S. Ave.	Sub-Group*	ACE Q-Score Percentile Rank					
		1-10	11-20	41-50	51-60	81-90	91-100
Below 1.60	1	13	15	18	19	23	26
	2	6	8	10	10	13	15
	3	21	24	29	30	35	38
	4	8	10	12	13	16	19
1.60-1.79	1	14	16	20	21	25	28
	2	7	8	11	11	14	16
	3	23	26	30	32	37	40
	4	9	11	14	14	18	20
2.40-2.59	1	19	22	26	28	32	36
	2	10	12	15	16	20	22
	3	30	33	38	40	45	48
	4	13	15	19	20	24	27
2.60-2.79	1	21	24	28	29	34	38
	2	11	13	16	17	21	24
	3	32	35	40	42	47	50
	4	14	17	20	21	26	28
3.60-3.79	1	29	33	38	39	45	48
	2	17	20	24	25	30	33
	3	42	45	51	52	58	61
	4	21	24	29	30	35	38
3.80-4.00	1	31	35	40	41	47	50
	2	19	22	26	27	31	35
	3	44	48	53	54	60	63
	4	23	26	31	32	37	40

* 1 = Veteran, junior college
2 = Non-veteran, junior college
3 = Veteran, other college
4 = Non-veteran, other college

TABLE V

CHANCES IN 100 OF GRADUATING IN ENGINEERING FROM FIRST QUARTER GRADE POINT AVERAGES

1st Quarter Average	Subgroup	Chances in 100
0.00-0.19	1	11
	2	7
	3	16
	4	9
0.20-0.39	1	12
	2	8
	3	18
	4	10

1.80-1.99	1	32
	2	25
	3	42
	4	28
2.00-2.19	1	35
	2	28
	3	46
	4	31

3.60-3.79	1	63
	2	54
	3	72
	4	58
3.80-4.00	1	66
	2	58
	3	75
	4	62

2. Non-veteran students from junior colleges (k = 61)
3. Veteran students from other colleges (k = 414)
4. Non-veteran students from other colleges (k = 122)

Examination of Table I readily reveals the fact that the four subgroups tended to respond to the criterion differently. The stratification was therefore maintained throughout the statistical analysis.

Since the sample had been stratified into four subgroups, to determine the relationship between the tendency to graduate in engineering and each of the prediction variables, within subgroups biserial coefficients of correlation were computed and are listed in Table II. All relationships were found to be significantly different from zero.

Modified discriminant functions using various combinations of the prediction variables were computed for within the subgroups and yielded multiple biserial coefficients of correlation indicative of the predictive ability of each scheme. It was immediately discovered that the linguistic score could be dropped from any prediction combination without incurring an appreciable loss in ability to predict. In Table III are shown the various combinations of prediction variables and the resulting multiple biserial coefficients of correlation. The correlations are indicative of the ability to predict for any of the four designated subgroups since, in the case of any of the foregoing discriminant functions, the coefficients ($a_1, a_2, \ldots, a_m$) to be multiplied by the appropriate prediction variable scores were the same for each subgroups. However, the constant (C) changed with each subgroup so that the predictions vary from one subgroup to another.

To facilitate the application of the discriminant function technique for such persons as counselors of students, tables of chances in 100 of graduating were computed by repeatedly solving a discriminant function. By means of a single table or a series of related tables a counselor can rapidly determine a transfer student's probability of graduating in engineering by entering the table or tables with the student's prediction variable scores. If the prediction were to be made on the basis of two prediction variables, a single table such as Table IV would be sufficient. For three variables, two tables are necessary. The number of probability tables to be consulted is one less than the number of variables to be used in the prediction when more than one prediction variable is involved.

If, prior to the date that a transfer student began his studies at the Iowa State College, it were desirable to predict his probability of graduating in engineering on the basis of his quantitative score and high school grade point average, a counselor could easily do so by means of a table such as that partially shown in Table IV. Because the quantitative scores are usually available as percentile ranks, the table was constructed accordingly.

The probabilities were computed for each of the four subgroups, ranging from 6 chances in 100 for the least qualified students to 63 chances in 100 for the best qualified students, and varying appreciably from subgroup to subgroup. In using the table the counselor would classify a transfer student in the appropriate subgroup, read up and across in accordance with the student's scores, and note his chances in 100 of graduating in engineering. This information based upon the academic success of an earlier group of transfer students with similar qualifications, certainly would provide the incoming transfer student with a more realistic basis on which to evaluate his position.

If a counselor were willing to wait until the completion of a transfer student's first quarter of studies at the Iowa State College before predicting the student's probability of survival, a table such as the one shown in part in Table V could be used. Table V resembles closely Table IV except that the prediction is now being made exclusively on the basis of the first quarter grade point average. On the basis of this variable, the probabilities of graduating ranged from 7 chances in 100 to 75 chances in 100. Again the probabilities differed among the subgroups.

The application of the modified discriminant function technique to the foregoing problem is but one of many applications. The prediction of the probability of attending college or of passing a specific course (5) are examples of prediction problems in which the investigator may wish to assume that the dichtomous criterion is in reality a single normally distributed variable. If, as in the case of the prediction of the probability of graduation, the ability to predict is found to be sizable, certainly the application of the technique to these and similar problems provides the instructor, the administrator, and the counselor with still another valuable tool.

REFERENCES

1. Ahmann, J. Stanley. Prediction of Achievement of Iowa State College Engineering Students Having Transferred from Other Institutions. Unpublished Ph. D. Thesis, Iowa State College, 1951.
2. Fisher, R. A. "The Use of Multiple Measurements in Taxonomic Problems," Annals of Eugenics, VII (1936), pp. 179-88.
3. Gowan, A. M. Unique Characteristics of

Freshman Veterans at the Iowa State College with Administrative Implications. Unpublished Ph. D. Thesis, Iowa State College, 1947.

4. Moore, J. E. "A Decade of Attempts to Predict Scholastic Success in Engineering Schools," Occupations, XXVIII (November 1949), pp. 92-6.

5. Sprain, Wilbur. Forecasting Probability of First Quarter Mortality in General Chemistry at Iowa State College. Unpublished M. S. Thesis, Iowa State College, 1951.

6. Wert, James E., and others. Statistical Methods in Education and Psychology (New York: Appleton-Century-Crofts Co., 1953).

CHARACTERISTICS OF THE EFFECTIVE TEACHER BASED ON PUPIL EVALUATIONS*

PERCIVAL M. SYMONDS
Teachers College
Columbia University

Part I

IT HAS long been recognized that the basic criterion of teacher effectiveness is the changes in pupils. Barr's[1] studies tested out teacher characteristics against the criterion of pupil achievement as measured by standardized tests, but most educators today believe that achievement tests measure only a small portion of the changes that we expect in children as a result of education in school. Research workers have been stymied in finding adequate measures of personality that can be given to groups inexpensively and scored or evaluated objectively.

It occurred to the writer that one of the important outcomes of education is the formation of attitudes by pupils, particularly attitudes toward the school, toward learning and toward teachers. And it should be possible to secure verbal expressions of these attitudes that have validity and such verbal expressions would represent an important outcome of education as well as a criterion against which to test teacher characteristics.

A form was devised to secure pupils' reactions to the following questions:

1. Which of your teachers makes the work most interesting?
2. Which of your teachers understands you best and likes you most?
3. Which of your teachers would you most like to have again next year?
4. Which of your teachers makes you most willing to study or to participate in the work of the class?
5. Which of your teachers helps you most to learn?
6. Which of your teachers would you feel most like consulting on some personal matter for guidance and counseling?
7. Which of your teachers most makes you want to continue to go through high school?

It will be noted that these questions do not ask a pupil to describe or to characterize teachers, but to react to them. In previous studies pupils have been asked directly to describe the teacher or teachers that they like best, and these descriptions have been taken as providing the qualities of the good or the effective teacher. Or pupils have been asked to rate teachers without describing them.

In this study pupils are asked to react to their teachers in answer to the seven questions given above. After the teachers had been placed on a scale according to the pupils' reactions or evaluations, they could be observed directly by trained observers in order to ascertain the qualities which differentiate those that are reacted to favorably from those that are reacted to unfavorably.

It will be noted that two of the items are double barreled in that two questions are asked in one. It is admitted that in question 2 teacher understanding and teacher liking are not identical, but the intention was to have pupils make a judgment as to which teacher they believed or felt reacted to them most favorably and which most unfavorably. Actually, this is a question for which pupils really do not have the evidence, so if they are to answer it at all, it must be on the basis of their feelings and judgment about the matter. Likewise, it is recognized that in question 4 studying and participating in the work of the class are not identical, but again the intention was to have pupils respond to their be-

* The research reported in this paper was supported by a grant from the Council on Research in the Social Sciences of Columbia University. Grateful acknowledgement for assistance in all stages of the study is given to Mr. Paul R. Robbins. Mr. Norman Gorin also assisted in the tabulations and computations, and Mr. Richard H. Bloomer participated in the observations and prepared one of the observation reports. Special credit is due to the principal of the school in which the study was conducted. She took an active interest in the study throughout, helped in making the arrangements for securing the pupil responses and made independent ratings of the effectiveness of the teachers in the school.

1. A. S. Barr. The Measurement of Teaching Ability (Madison, Wis.: University of Wisconsin Press, 1946).

liefs or feelings as to which teacher stimulated them to the greatest activity.

A ranking method was used rather than a rating method as it is believed that ranking forces a discrimination that one does not obtain from rating.

As the form was drawn up, pupils were first asked to write the names of all the teachers with whom they have classes this year. This became a master list to serve later in the rankings of their teachers to the seven questions.

On the average a pupil had six teachers, but in some instances a pupil had fewer than six teachers and in some instances as many as seven or eight. Three or four of these are teachers of the so-called "prepared subjects." The rest are teachers of special or elective subjects such as physical education, art, music, foods, metal work, wood work, printing.

Name, grade and section were not asked for until the end of the blank so that pupils filling it out would not be encouraged to govern their responses so as not to offend anyone, as they would have been if they had written their names at the beginning of the exercise. They were also asked to write the name of their homeroom teacher at the end, partly for purposes of identification.

This method of determining the characteristics of the effective teacher was tried out in a junior high school in one of the suburbs of New York City. The community has a large residential as well as a large industrial area. The school serves a lower middle class residential area. It is coeducational and includes a substantial Negro as well as white population. The school is staffed by 32 teachers, and 453 pupils filled out the form, distributed as follows:

TABLE I

DISTRIBUTION OF PUPILS WHO TOOK PART IN THE STUDY BY GRADE AND SEX *

	Number of Pupils Filling Out the Form		
	Boys	Girls	Total
7th Grade	61	72	133
8th Grade	66	80	146
9th Grade	91	83	174
Total	218	235	453

* This tabulation includes only those who were present on the day of the study and who came to the auditorium.

The forms were filled out in the school's auditorium by three groups of pupils representing the three grades in the school. At a given time all of the classes of a given grade were brought into the auditorium; and after they were seated, their teachers were excused and were not present while the pupils made their rankings. The only adults present were the investigator, his assistant and the principal of the school.

After an explanation to the pupils as to why they had been assembled, the blanks were distributed. Pupils were paced so that they might be kept together as is done in group testing. They were first requested to write the name of their teachers on the lines provided for this purpose at the top of the first page. Then the first question was read aloud, and they were instructed to fill in on the lines the names of their teachers in order of making the work interesting. In case they had work with more than six teachers, they were instructed that they might write names below the last line but that they were to make sure that the teacher's name which stood at the bottom of the list was the one who made the work least interesting. When all or practically all of the children had finished, the second question was read to them, and they proceeded with the ranking of teachers to the second question.

In general, few questions were asked. If the procedure were to be repeated, it is believed that it would be better if every pupil could have in his hands the names of all the teachers in the school, for it was discovered that some pupils even in November did not know the names of some of their teachers, and many more did not know how to spell the names that they knew.

Pupils were cautioned not to include the names of their homeroom teachers in their rankings unless their homeroom teacher also served as a teacher in one of their classes. They were also cautioned to list only those teachers with whom they had classes this year. In some instances there was a great temptation to include the names of homeroom teachers or teachers from previous years who were very much liked.

Surprisingly enough the staff of the school appeared to show no interest in what was going on and apparently never learned the nature of the exercise. It was conducted so smoothly that pupils did not comment on it to their teachers, and apparently auditorium exercises came so frequently that teachers had little curiosity about it.

It was the principal's decision to conduct the study without consulting the faculty or taking its members into her confidence or asking their judgment. Other principals might not want to conduct a school activity of this kind without democratically letting the faculty in on it. But the principal of this school, and the staff also, felt secure enough that it could be worked out without their knowing about it. Some half hearted attempts which were made by members of the staff to ask the principal about it were easily thrown off by the principal as "just some research by a man from Columbia."

If this method were to be used in a practical situation the reactions of the staff to the study could not be disposed of so easily, and many problems of securing the cooperation of the staff without their interference would have to be worked out. But fortunately such problems did not come up in this study.

The main task in this part of the study was to find the mean ranking of each teacher for each of the seven questions, as well as the average mean rank. This meant the computation of 256 means. First, however, the rankings for each teacher for each question by each pupil had to be tabulated. In making these tabulations the pupils were separated by grade and by homeroom so that it would be possible to make separate means by grade, sex of pupil and by homeroom teacher.

What to do when the number of teachers for any one pupil failed to reach or exceeded six presented a technical difficulty which had to be overcome by a special device. It was decided to transmute the ranks given by these pupils into equivalent ranks in which the highest number was 6 and with the assumption that the variances of the different sets of ranks should be equal[2]. These transformations were made in accordance with Table II.

The mean rank for each teacher for each question was further complicated by the fact that each teacher was ranked by a different number of pupils. For instance, one teacher of music was ranked by 324 pupils on question 2 with an average rank of 3.86 while a part-time teacher of music was ranked by 10 pupils with an average rank of 2.15. It was found that the correlation between the mean ranking and number of pupils used in determining a mean was +.336 (based on question 7) - that is, there is a tendency for teachers with the larger pupil load to have higher ranks (to be ranked lower on the list). It is believed, however, that this is a real tendency and does not invalidate the method of comparing pupils' evaluation of their teachers by the ranking method.

Table III gives the mean ranks for the 32 teachers in the school in which the study was conducted and the average of these mean ranks.

One of the goals of the study was the determination of the relative position of the teachers in the school as judged by the pupils' answers to the seven questions. In ranking their teachers the pupils were not given a choice to tell how interesting they found the work of the school, how much they learned or whether they would like to continue on through high school on any absolute scale. Rather they were forced to discriminate between their teachers in the extent to which one teacher helped them reach these subjective states more than another.

It is believed then that the mean rank serves as a measure of a teacher's effectiveness as represented by the seven questions. A low mean rank means that pupils thought highly of that teacher; while a high mean rank means that pupils placed that teacher below his or her colleagues on the question.

Teacher 9 received a mean rank of 1.45 on question 7; while teacher 15 received a mean rank of 4.81. That means that many pupils gave teacher 9 first and second rank; while many pupils gave teacher 15 fourth, fifth or sixth rank. These differences are graphically illustrated by showing the actual distribution of ranks - Table IV.

The distribution of means for each question is given in Table V.

It may be noted that the mean of the mean ranks falls below the midpoint of the ranks-3.5. This is due to the fact that there is a correlation between mean rank and number of pupils who contributed to each mean. Since teachers who have many pupils tend to have higher (higher numbered) ranks, there are more teachers with lower (lower numbered) ranks which brings the average of ranks down.

It may also be noted that some of the distributions of mean ranks have larger variances than others. (See Table III.) This is believed to indicate a difference in the discrimination with which certain questions are answered. In some questions there is apparently greater pupil agreement and hence mean ranks of high and low teachers can move more closely toward 1 and 6. On this basis, pupils answered question 5 on "Which teacher helps you to learn the most?" with greatest discrimination, and question 2 on "Which teacher understands you best and likes you most?" with least discrimination.

The variation in the mean rankings for any one teacher may be in part merely a function of the number of pupils that ranked this teacher. The fewer the pupils, the greater the variation in means that may be expected. But some of these differences probably have real significance. For instance the variation in means for Teacher 27 cannot be explained on the basis of small N for about 110 pupils contributed to her means. It will be noted that this teacher had lower rankings (larger rank numbers) for being approachable for counseling than for helping to learn, and it is probable that this is a real distinction in the case of this teacher.

The odd-even coefficient of reliability computed for question 7 (to serve as a sample of all of the questions) was .873. Raised by the Spearman-Brown formula, the coefficient of reliability of determination of the means for question 7

2. This procedure was followed on the advice of Dr. Herbert Solomon, Teachers College, Columbia University, to whom the writer acknowledges his thanks.

TABLE II

FOR TRANSMUTING RANKS THAT CORRESPOND TO RANKS OF 1 TO 6 WITH EQUAL VARIANCES

	Number of Persons to be Ranked						
	2	3	4	5	6	7	8
Ranks Assigned	5	6	6	6	6	6	6
	2	3.5	4	4	5	5	5
		1	3	3.5	4	4	5
			1	3	3	3.5	4
				1	2	3	3
					1	2	2
						1	2
							1

TABLE III

MEAN RANK OF 32 TEACHERS FOR EACH OF 7 QUESTIONS AND THE AVERAGE OF THESE MEANS

Teacher	Question Number 1	2	3	4	5	6	7	Average
1	2.72	2.15	2.50	3.28	3.56	2.35	2.93	2.78
2	4.26	3.86	4.29	4.13	4.77	4.03	4.19	4.22
3	3.40	3.47	3.72	3.89	4.08	3.54	4.00	3.73
4	2.75	2.72	2.80	2.99	2.67	2.37	2.40	2.67
5	4.37	4.05	3.82	4.39	4.65	3.87	4.25	4.20
6	3.65	3.72	3.50	3.49	3.53	3.42	3.22	3.50
7	3.24	2.84	3.57	3.15	2.81	2.47	2.86	2.99
8	3.71	3.82	4.06	3.72	3.93	3.44	3.62	3.75
9	1.62	1.67	2.00	1.24	1.15	1.69	1.45	1.54
10	2.32	2.51	2.20	2.21	2.22	3.08	2.19	2.39
11	3.85	3.36	3.33	3.70	3.68	3.13	3.86	3.56
12	3.38	3.79	3.20	3.49	3.53	3.92	3.74	3.58
13	3.63	3.85	3.63	3.96	3.64	4.24	4.24	3.88
14	3.76	3.38	3.34	3.05	4.20	3.56	3.06	3.48
15	4.48	4.49	4.20	4.38	4.23	3.93	4.78	4.35
16	2.41	2.38	2.78	2.17	1.54	1.74	2.10	2.16
17	3.39	3.20	3.05	2.95	2.59	2.52	2.66	2.91
18	4.03	3.48	3.75	3.99	3.48	3.30	3.78	3.69
19	4.02	3.46	3.33	3.63	3.73	3.61	3.37	3.59
20	3.21	3.26	3.50	3.55	3.58	3.61	3.52	3.46
21	2.79	2.92	2.85	3.73	3.37	3.17	3.01	3.05
22	3.93	3.64	3.85	3.96	4.05	3.85	3.94	3.89
23	2.22	3.34	3.50	2.09	3.51	3.45	3.38	3.07
24	3.12	3.38	3.33	3.32	3.45	3.73	3.31	3.38
25	3.62	3.86	3.94	3.52	2.71	3.78	3.59	3.57
26	2.90	2.35	3.05	2.65	2.24	2.76	2.37	2.61
27	3.65	3.65	4.53	3.52	2.78	4.74	4.34	3.88
28	1.52	1.89	1.31	1.98	1.86	1.84	1.89	1.76
29	4.13	4.44	4.48	4.44	4.78	4.44	4.56	4.46
30	2.30	2.56	2.36	2.36	2.07	3.12	2.40	2.45
31	2.61	2.42	2.26	2.26	2.36	2.68	2.55	2.44
32	3.55	4.41	3.56	3.85	4.12	2.93	3.62	3.72
M	3.27	3.26	3.30	3.27	3.28	3.26	3.29	3.27
σ	17.1	16.3	16.4	17.8	20.8	17.0	19.0	16.0

TABLE IV

DISTRIBUTION OF RANKS FOR TWO TEACHERS ON QUESTION 7

	Teacher 9				Teacher 15		
	Grade				Grade		
Rank	7	8	9	Total	7	8	Total
1	7	22	1	30	3	1	4
2	4	1		5	4		4
3		1		1			
3.5	1			1		1	1
4	1			1	3		3
5					1	2	3
6	1			1	9	16	25
Totals	14	24	1	39	20	20	40
	Mean Rank 1.45				Mean Rank 4.81		

TABLE V

DISTRIBUTION OF MEANS OF RANKS FOR EACH QUESTION

	Question							
Mean Rank	1	2	3	4	5	6	7	Average
4.75-4.99	-	-	-	-	2	-	1	-
4.50-4.74	-	-	1	-	1	1	1	-
4.25-4.49	3	3	2	3	-	1	2	2
4.00-4.24	3	1	2	1	5	2	3	2
3.75-3.99	3	5	4	5	1	5	3	4
3.50-3.74	6	3	7	6	8	5	5	8
3.25-3.49	3	8	4	4	3	4	3	3
3.00-3.24	3	1	3	3	-	4	3	2
2.75-2.99	3	2	3	2	2	2	2	3
2.50-2.74	2	3	1	1	3	2	2	2
2.25-2.49	3	3	2	2	1	3	3	3
2.00-2.24	1	1	2	3	3	-	2	1
1.75-1.99	-	1	-	1	1	1	1	1
1.50-1.74	2	1	-	-	1	2	-	1
1.25-1.49	-	-	1	-	-	-	1	-
1.00-1.24	-	-	-	1	1	-	-	-
N	32	32	32	32	32	32	32	32

becomes .933. This is a substantially high coefficient of reliability and indicates that the evaluation of teacher effectiveness by pupil ranking is a highly reliable affair.

Determinations of the mean rank given by pupils to each teacher were made by grade and also by boys and girls taken separately for question 1.

The rank-difference correlation between the mean ranks given by grades were as follows:

r Grade 7 - Grade 8 = .752
r Grade 8 - Grade 9 = .621
r Grade 8 - Grade 9 = .413

There is no obvious explanation as to why there should be less agreement in sizing up teachers on question 1 between grades 8 and 9 than between 7 and 9, and 7 and 8. These correlations are low enough to indicate that there are real differences in the appraisal of teachers from grade to grade. The principal of the school corroborated some of these differences from her own observations. In the case of one mathematics teacher, she reports that seventh grade pupils meeting her for the first time fear her, but by the time pupils have reached the eighth grade they understand her, are no longer afraid of her and like her. Another teacher is described as being bossy and dictatorial with younger pupils but treats older girls as "pals." It would seem as though a different set of factors was required to explain these differences in each case.

The rank-difference correlation between the mean ranks given by boys and girls on question 1 was .522. This correlation is low enough so that it would appear that some teachers are actually responded to more favorably by boys than girls and vice versa. With regard to one teacher, the principal states, "He is very shy of girls and ill-at-ease with them. He was not able to control a class of girls which he had taught for one semester although his control of boys' classes and his rapport with the boys is excellent." In the case of another teacher, "The girls respect her because she is a 'lady', whereas the boys are downright unappreciative of her efforts to instill some culture into their noisy personalities." So again in the case of the different responses of boys and girls the factors are somewhat different in each case.

The correlation of each of the questions with the others should provide information concerning their relative independence. Table VI presents these intercorrelations.

It should first be noted that these correlations are high. The lowest in the table is .697 between questions 5 and 6 (Which teacher helps you to learn the most? and Which of your teachers would you feel most like consulting?), whereas questions 4 and 7 correlate .915 (Which teacher makes you most willing to study? and Which teacher makes you most want to continue to go through high school?). The mean of these 21 correlations is .821.

It should be evident that whatever discrimination there is in answering the questions is offset by a very large halo effect. Basically pupils rank their teachers in terms of some underlying liking or feeling which tends to influence the rankings to each question.

The line at the bottom of the table headed Average is simply the average of the correlations of each of the questions with the other 6. The final column (8) headed Mean of Means is the correlation of the mean rank for a question with the mean of the mean rank for all questions. The correlations in this last column are extremely high, five of them being over .90. There is an artifact here because the mean rank of each question is included in the mean of the means of all questions. Even so these correlations are so high as to indicate that question 7 and possibly question 4 provide about the same information concerning the relative positions of the teachers as does the average of them all.

In order to obtain evidence concerning the validity of the rankings by pupils, the principal of the school was asked to rate the teachers on three characteristics: Disciplinary Control, Relationship with Pupils, and Success in Terms of Pupil Achievement. In addition she was asked to rate each teacher on All-round Excellence. In preparing ratings the principal consulted with the assistant principal. A sample of the rating scale used is given at the top of the next page.

It is believed that this rating scale is admirably adapted to the purpose for which it was intended. All the teachers' names are included on the rating sheet so that it was possible to compare the ratings given to each teacher and to make the necessary discriminations. The ratings were made on a six-point scale which comes close to the optimum for rating human characteristics.[3] An even number of steps in the rating scale favors a discrimination between the two middle positions. A rough guide in terms of percents is given for the number which should go in each column. Roughly this means that two or three should be marked in columns 1 and 6; five or six in columns 2 and 5; and the remaning sixteen divided equally between columns 3 and 4. It was intended that this be a rough guide only; but following this suggestion insures that the whole range of the scale be used instead of only a portion of it as is so often the case. The suggestion was given that the checks for columns 1

3. P. M. Symonds. "On the Loss of Reliability Due to Coarseness of the Scale," Journal of Experimental Psychology, VII (1924), pp. 456-60.

TEACHER RATING SCALE

Rate on a scale from 1 to 6 as follows:

6 Very superior	(7%)
5 Superior	(18%)
4 Above average	(25%)
3 Below average	(25%)
2 Inferior	(18%)
1 Very inferior	(7%)

Trait to be rated _____

Place a check in the appropriate column

Teacher	1	2	3	4	5	6	

and 6 first be made; then for columns 2 and 5 next and finally those for columns 3 and 4 last.

Table VII gives the intercorrelation between these four sets of ratings as well as with the average of the four.

These correlations are substantially high. The average of the correlations of discipline, relationship with pupils and pupil achievement is .603 The average of the correlations withall-round excellence is .731. Apparently the principal of the school introduced considerable discrimination in rating these different characteristics so that the influence of the halo effect is not pronounced. The correlations with the averages of all four ratings (again this contains an artifact due to overlapping) is in the high .80's, and the correlation of this average of all four ratings with the rating for all round excellence is .916. It is believed on the one hand that the principal of this school knows her teachers well and on the other hand, that she used considerable care in making these ratings.

Finally Table VIII presents the correlations of the pupil rankings with the principal's ratings.

An explanation of the minus signs in Table VIII is first in order. A lower number in pupil ranking (1 or 2) signifies excellence whereas a high number in principal rating (5 or 6) signifies excellence. A negative correlation indicates, therefore, that pupils and the principal tend to agree in their evaluations.

It may first be noted that the correlations of the principal's ratings with the pupils rankings are monotonously even. The correlations of discipline with the rankings ranges only between -.57 and -.69; of principal's ratings of relationship with pupils between -.56 and -.78; of principal's ratings of teacher success in eliciting pupil achievement, between -.40 and -.48; and of principal's ratings of all round excellence between -.45 and -.57.

Actually the correlations of principal's ratings of "teachers' discipline" with "relationships with pupils" are about the same as the correlations of pupils' ranking with these principal's ratings. We would have to conclude that since the pupils agree with principal's ratings of teachers' discipline as well as the principal agrees with herself in rating two different characteristics, the validity of the rankings by the pupils cannot be much less than validity of the ratings by the principal.

These correlations of pupil ranking with principal rating of teachers is higher than is customarily found. Hedlund[4] reports correlations between supervisors and pupil evaluation of teacher effectiveness, using two types of appraisal. First, special instruments were constructed to be used by supervisors and pupils. The instrument for use by the supervisors consists of 19 items of "critical behavior" the presence or absence of which is to be noted, while the instrument to be used by pupils consists of 33 items of "critical behavior" to be similarly answered. Second, both principal and pupils gave overall judgments.

4. R. H. Hedlund. The Cooperative Study to Predict Effectiveness in Secondary School Teaching, Third Progress Report (New York: New York University, 1953).

TABLE VI

INTERCORRELATIONS OF THE MEAN RANKS OF THE SEVEN QUESTIONS

Question	Question							Mean of Means
	1	2	3	4	5	6	7	
1		.840	.819	.912	.799	.734	.838	.911
2			.840	.848	.821	.813	.891	.906
3				.799	.729	.760	.853	.878
4					.874	.753	.915	.943
5						.697	.856	.903
6							.860	.868
7								.961
Average	.823	.842	.800	.850	.795	.769	.869	

TABLE VII

INTERCORRELATIONS OF THE PRINCIPAL'S RATINGS

	Discipline	Relationship with Pupils	Pupil Achievement	All-Round Excellence	Average
Discipline		.669	.576	.683	.866
Relationship with Pupils			.563	.789	.878
Pupil Achievement				.720	.831
All-Round Excellence					.916
Average	.643	.674	.620	.731	

TABLE VIII

CORRELATION OF PUPIL RANKINGS WITH PRINCIPAL'S RATINGS

Question	Discipline	Relationship with Pupils	Pupil Achievement	All Round Excellence	Average
1	-.651	-.691	-.397	-.560	-.646
2	-.598	-.735	-.446	-.531	-.668
3	-.595	-.721	-.480	-.524	-.677
4	-.625	-.744	-.426	-.522	-.674
5	-.695	-.561	-.411	-.451	-.600
6	-.678	-.781	-.446	-.571	-.721
7	-.573	-.758	-.416	-.561	-.673
Average	-.678	-.774	-.479	-.589	-.675

The correlations which he found are given in Table IX.

TABLE IX

CORRELATIONS BETWEEN PRINCIPALS' AND PUPILS' RATINGS IN COOPERATIVE STUDY TO PREDICT EFFECTIVENESS IN SECONDARY SCHOOL TEACHING

	Score based on pupils' instru.	Pupils' overall judgment
Score based on super'ors instru.	.38	.26
Super'ors' overall judgment	.44	.31

These data were collected with great care, but represent typical conditions covering many schools. It is believed that the present study yields substantially higher correlations for two reasons: (1) The method of ranking by which the pupils indicated their judgment probably yielded results with very high reliability, and (2) the principal took great pains in making her ratings. She knows her staff intimately, and there is reason to believe that she has good judgment in assessing the work of individual teachers.

Looking at the correlations of pupils' with principal's ratings more closely, it appears that the principal's ratings of teachers' discipline correlates most sharply with pupils' rankings on "Which teacher helps you to learn the most?" (-.678); principal's ratings of teachers' relationships with pupils correlates most sharply with pupils' rankings on "Which teacher would you feel most like consulting?" (-.781); principal's ratings on teachers' ability to elicit pupil achievement correlates most sharply with pupils' rankings on "Which teacher would you most like to have next year?" (-.480); principal's ratings on teachers' all-round excellence correlates most sharply with pupil ranking on "Which teacher would you feel most like consulting?" (-.571). There is not too much rhyme or reason to these highest correlations and we must conclude that since the correlations with any one of the principal's ratings are so much alike the halo effect plays a very important part in the pupils' rankings, and the differences in the rankings by pupils of any one teacher on the several questions are not too significant.

One final set of correlations shows the relation with another set of ratings made by the principal on the teachers in the school later in the year. The ratings already reported are relative ratings—that is, the ratings on a six point scale were done in such a way that more attention was given to above and below average ratings for the group and to the highest and lowest rating than to the descriptive titles superior or inferior. The new set of ratings to be reported are absolute in the sense that each rating indicates a well defined judgment on the part of the principal about each teacher.

The following "Scale of Teacher Effectiveness" indicates the nature of the second rating by the principal.

Scale of Teacher Effectiveness

4. An outstanding teacher, probably destined for a more important position, either administrative or on a higher teaching level.

3. A satisfactory and acceptable member of the school faculty. A good teacher, good relationships with pupils, cooperative in school affairs, and in general a helpful member of the team. Most of my teachers will fall into this category.

2. Possesses limitations as a teacher and as a member of the staff. If I had free choice, I believe it would be possible to find another person who would be more effective in his (or her) place.

1. In my judgment this person is so ineffective that he (or she) should no longer serve as a teacher.

As might be expected, the ratings were highly concentrated at 3. Consequently, the variance was restricted as will be seen from Table X.

TABLE X

DISTRIBUTION OF PRINCIPAL'S RATINGS ON SCALE OF TEACHER EFFECTIVENESS

Scale Value	No. of Teachers
4	1
3	22
2	7
1	2
	32

The correlations of the principals's ratings on the Scale of Teacher Effectiveness with the earlier ratings on the 6-point scale are given in Table XI.

TABLE XI

COEFFICIENTS OF CORRELATION BETWEEN PRINCIPAL'S RATINGS ON SCALE OF TEACHER EFFECTIVENESS WITH EARLIER RATINGS

Disciplinary Control	.508
Relationships with Pupils	.588
Success in Terms of Pupil Achievement	.386
All-Round Excellence	.649
Average	.548

It is of considerable interest that in the last analysis the principal rates her teachers more on relationship with pupils than on the ability of the teacher to produce high achievement.

The correlations of the principal's second rating of teachers with the rankings of the teachers by the pupils is given in Table XII.

TABLE XII

COEFFICIENTS OF CORRELATION BETWEEN PRINCIPAL'S RATINGS ON THE SCALE OF TEACHER EFFECTIVENESS AND PUPIL RANKINGS OF TEACHERS

Question No.	
1	-.426
2	-.478
3	-.397
4	-.396
5	-.306
6	-.460
7	-.422
Average	-.458

The correlations of the rankings by pupils with the second rating by the principal are not much lower than the correlations with the principal's earlier ratings. This is further evidence of the essential validity of the pupils' rankings using the principal's ratings as a criterion.

The tabulation of the pupils' rankings was laborious and time consuming and open to certain inaccuracies due to the difficulty in some instances of deciphering pupils' handwriting or making decisions when pupils failed to follow instructions exactly. The possibility of using pupils' expression of the best and worst teacher in answer to each question was considered as an alternative procedure, a procedure which would greatly cut down time and labor and which would reduce inaccuracies.

Question 7 was retabulated twice, once using only the highest and lowest rankings, giving their values of 1 and 6 respectively; and also the two top and the two bottom rankings, giving values of 1 and 2 and 5 and 6 respectively. The rank-difference correlation of the average rankings using only the two top (1 and 2) and two bottom rankings (5 and 6) with the average rankings using all the rankings was .976. Using only the top and bottom rankings, the correlation with the method using the complete rankings was .936. This latter figure is high enough so that for all practical purposes one can secure adequately accurate results by simply asking pupils to name their best and poorest teacher to any question that one may wish to ask pupils about their teachers.

As a practical procedure then, one can secure reliable and valid estimates of the effectiveness of teachers by asking the pupils in any school, grade or class to simply indicate which of their teachers they would place at the top and at the bottom in response to the kinds of questions used in this study. If those teachers who are mentioned as "best" are given a value of 1 and those "poorest" a value of 2 and an average of these values found, the result will be a valid and reliable measure of that teacher with regard to the pupils' estimate of her on the given question. The better teachers will have values closer to 1, and the poorer teachers will have values closer to 2.

Illustration

Teacher A

Rated "best" by 115 pupils	115 × 1 =	115
Rated "poorest" by 6 pupils	6 × 2 =	12
	121	127
	Divided by	121
	Average Place	1.05

Teacher B

Rated "best" by 12 pupils	12 × 1 =	12
Rated "poorest" by 67 pupils	67 × 2 =	134
	79	146
	Divided by	79
	Average Place	1.85

Part II

The first part of this study just described tells how 32 teachers in a junior high school were ranked by their pupils and rated by the principal of the school on a number of characteristics. This provides us with information about these teachers and makes it possible to characterize

each teacher as more or less effective from the point of view of the pupils and the principal.

In the second part of the study, 17 of these teachers taken from the top and bottom of the list were observed while teaching. It was hoped that as a result of these observations it should be possible to isolate those characteristics which distinguished the better from the poorer teachers.

The method of making these comparisons was subjective inasmuch as this was a "finding" and "exploratory" study rather than a "definitive" study. Now that the spade work of this study has been completed, the basis for making a definitive study has been laid.

A few comments on method are in order. In conducting an observation the observer would try to enter the room before a period began, would introduce himself to the teacher and ask his permission to visit the class. He would state that he would like to sit unobtrusively in the rear of the room and would appreciate it if the teacher would conduct the lesson as planned without any reference to his presence.

It is the writer's practice to take notes as he observes. There are many who object to this procedure for fear of embarrassing the teacher. In the writer's experience he has seen no evidence that note taking interferes in any way with the teacher's methods or procedures. Usually the teacher becomes absorbed in the work of the hour although there is evidence that most teachers are well aware that a visitor is in the room. Those who believe that the teacher is sensitive about note taking are at any rate sensitive themselves and in most cases that is as far as it goes.

The notes as taken by the writer are in the form of a running commentary. They describe what the teacher does and pupil response. They contain little phrases that the teacher uses or brief snatches of conversation. They include impressions and comméntaries that the observer wished to record. They include characterization of the teacher, his attitude toward pupils and toward the work.

After the observation is over, the notes are reviewed and reorganized and a report of the results of the observation is then written.

During the observation the writer is continually asking himself "Why is this teacher considered a good (or poor) teacher by the pupils or the principal? What is there about this teacher that makes him stand high (or low) in the regard of his pupils?"

Reports which were written after observing a teacher have been analyzed by listing each comment descriptive of teacher behavior or attitude. Such a listing is not altogether satisfactory, inasmuch as the same behavior or attitude would be described by synonymous terms or by terms with slightly different meanings. In addition there were many terms and phrases which expressed slightly different variations or aspects of a given theme.

Finally it appeared that the two extreme groups of teachers could be differentiated on three clear cut bases.

1. The superior teachers liked children; and the inferior teachers disliked children.
2. The superior teachers were personally secure and self-assured; while the inferior teachers were personally insecure and had feelings of inferiority and inadequacy.
3. The superior teachers were well integrated and possessed good personality organization; while the inferior teachers were personally disorganized.

The Superior Teacher: Liking Children

With regard to the first point, the superior teachers showed their liking for boys and girls in many ways. They were interested in their pupils, knew their names, knew something of the pupils' interests, pecularities and backgrounds. In general they saw each pupil as someone with possibilities for growth and development. Some pupils learn rapidly, but other pupils require patience. Pupils who were deviates were accepted with tolerance.

> During one of her unassigned periods, Miss P., a language teacher talked animatedly about her work and about her plans for the next period. She spoke about individual children in her class with some detail and sympathetically. One boy comes from a difficult home. The mother has hailed the father into court, and the boy comes to school disheveled and unable to concentrate. One girl has come recently from Czechoslovakia and is still mastering the English language. One Negro girl has recently come from the South, and this is the first school that she has attended with white pupils. Another Negro girl has a low I.Q. These descriptions of individual children indicated how interested she was in them and how much she cared about them.
>
> Miss I. feels responsible for each pupil, gets to know him and takes a personal interest in his progress. Boys whom other teachers find very difficult seem to appreciate Miss I.'s interest in them and become very loyal to her.

Each pupil was respected as an individual. If he had a question, he was listened to attentively, and whatever his request it was given consideration. These teachers were helpful in making explanations in terms and at a tempo so that a pupil could understand.

Miss P. goes about the room helping individuals who seem to be having difficulty planning their work. She seems to know each pupil well and what she can expect from him. The boy from the broken home is restless, annoys other pupils, and draws pictures. She tells me that one must have a sense of humor in dealing with a boy of that kind. To one of the girls who has difficulty settling down she said, "It is time you grew up."

In general the superior teachers respected and trusted their pupils and gave them responsibility for phases of the class work. It was observed that pupils who were given responsibility carried out their assignments faithfully to the best of their ability.

Pupils were quite orderly and respectful, and apparently were accustomed to working with good relations with Miss P. At the beginning of the hour, the class broke up into committees which had previously been formed in order to make plans for working up units in their project preliminary to going to the library. One girl was given the responsibility of going to the library to find out how many pupils could be accommodated at different periods of the day, and when she returned she was given the additional responsibility of going about and assigning different pupils to different library periods.

Mr. T., an English teacher, always referred to the class with the term we, so that comments like "Remember, what did we read about this last week?" occurred frequently. Several times hesitant children were encouraged with a phrase like, "Come on, you know it, don't be afraid to say it." And after a hesitant child has made a right answer, "There, you knew it all the time, Bill." Mr. T. invariably used a question to answer a question as if he were drawing out the responses and at the same time including the children in the category of people whose opinions were to be respected.

The superior teachers did not noticeably have favorites or give special consideration to individuals or groups. In general they made reasonable requests and assignments, and if special situations arose (as when a group of girls were planning to travel to another city in the evening to attend a social function) assignments were adjusted to the circumstances.

The superior teachers exhibited a certain warmth in their relations with pupils. Instead of maintaining distance they mingled with the pupils in a friendly fashion. They were agreeable and pleasant in mien and bearing. Occasionally they would show genuine affection by a pat on the shoulder or by the use on an endearing expression. They showed genuine appreciation of a child's work. In general they showed satisfaction if a pupil showed superior achievement, but they also showed genuine recognition when it was clear that a pupil had done all that could be expected of him under the circumstances. They challenged on the one hand and were encouraging on the other.

Mr. BB. plays fair with the boys in his woodworking class. They understand just where they stand with him. He believes that a boy likes to be held up to standards. On the other hand, he tends to be lenient in marking if he believes that there is reason for it. He says you have to know your boy. If a boy, for instance, does not have coordination but has tried, he will give him a mark which is somewhat higher than the number of errors calls for. He says that he sees no merit in hurting a boy, and to give him a higher mark than his work calls for may help to encourage him.

The Inferior Teacher: Disliking Children

The inferior teachers, on the other hand, showed clearly that they disliked their pupils. The inferior teachers sometimes did not know the names of their pupils or knew little about their interests or their lives outside of school.

Although Mr. J., a teacher of metal working has had his class for the whole half year he did not know a boy's name and has to ask him what his name is. He emphasizes how careless the boys are with their tools and materials, and how they need constant supervision.

One sure sign of this dislike was when they would tell the observer how difficult a certain pupil was, or how stupid, lazy or troublesome certain pupils might be.

Miss CC., an art teacher, told the observer that one boy who was separated from the class had a bad heredity, that his parents did not want him, that he has been picked up by the police for delinquency, and that she does not trust him. When her back is turned he is up to little pranks, such as turning on the water faucet or annoying other children.

All inferior teachers had a somewhat hopeless attitude toward some of their pupils and seemed to expect the worst from them. In several instances there were signs of bad temper when the teacher would show pronounced irritation at things that pupils were or were not doing. They tended to be very critical of individ-

ual pupils and their work and seemed never satisfied with what pupils might accomplish. They were quick to point out faults.

The inferior teachers did not respect individual pupils. They gave scant attention to their requests and questions, and took every opportunity to embarrass and to humiliate them. They indulged in sarcasm and ridicule as a method of control.

> A boy comes in from another class with a request from the teacher for a piece of wood for a certain purpose. Mr. J. questions the boy sharply with regard to the dimensions of the piece that he wants, and the other boys in the class find amusement at the discomfort of the boy who is being questioned.

The inferior teachers gave special consideration to certain pupils or to groups who were their favorites. For instance, some of the inferior teachers gave special consideration to the brighter pupils, and one seemed to favor white pupils. On the opposite side other pupils were specially singled out for censure and criticism.

> Miss E. was to some extent partial to the white girls in her classes. She seemed to know the names of most of them and used them frequently. During one class when the girls were nominating players for an all-school game, she acknowledged the hands of about twice as many white girls as Negroes, though the proportion of hands was about equal. In addition, she singled out only two white girls for her fixated gaze or sarcastic comment, while the total number of such incidents was about twenty-five.

The inferior teachers showed their dislike of and disinterest in pupils by their interest in materials and subject matter. An art teacher showed keen interest in color and form; while a teacher of wood-working showed interest in his tools and materials.

The inferior teachers were cold, distant, unapproachable. Whatever they did seemed to create a barrier between them and their pupils. A notice was read monotonously and rapidly so that pupils had difficulty in following; or an explanation was given carelessly or with vocabulary that pupils could not understand, or the explanations were long and complicated.

In some instances their complaining went beyond the individual pupils in their classes to embrace the administration, the school building and equipment, the curriculum and text books, in fact, any aspect of education.

> Miss B. a teacher of history, explained that this particular class was a very difficult group of nonreaders in the ninth grade. A week ago she gave the Iowa Every Pupil Test, and the scoring of the test pretty nearly ruined her. She said that she had had a headache over the weekend, the remains of which she still felt today.
>
> There is a considerable undertone of talking among the pupils. Apparently they pay very little attention to her, even while she is talking to them.

The Superior Teacher: Inner Strength

With regard to the second point, the superior teachers possessed a kind of inner strength and security which exhibited itself in many characteristic ways. The exact nature of what is meant by "inner strength" may be obscure and puzzling. In part, it is the absence of fear of one's own aggression, that is, one's assertiveness and firmness in dealing with others. In part, it is a belief in one's self and in what one stands for. The person with inner strength stands secure, has no need to belittle or humble himself, has no cause to doubt his status or his abilities.

So the superior teacher tends to be assertive and to exercise firm control. In the classes of the superior teachers visited, there was no doubt who was in control, and as a result the classes were usually orderly and well disciplined.

> Miss I. states that pupils do not mind if she is fussy but apparently like it. On the other hand, she takes pride in being precise. There is no fooling in her classes, and apparently pupils enjoy the challenge that she gives them and like being held rigidly to a high standard of perfection. The impression was gained that the pupils like to be held under control in this fashion when it is not done harshly or punitively.
>
> Mr. K. is an earnest, businesslike man, and the pupils evidently respect him because he works hard and conscientiously at the job. He controls the class partly by his strength of voice and his earnestness. A boy goes to the board. He says, "Watch Joe. If he makes an error, do not say anything until he is finished. Match your work with his." And to Joe he says, "Stay right there, Joe. I am not going to bite you." Mr. K. is liked by the boys and girls because of his forthright, honest, businesslike manner.

Actually, discipline seldom appeared as a problem - little attention was paid to it - so that the attention of everyone was focused on the work and the activities of the class. If a pupil became disorderly, it was not overlooked (if it was genuine disorder), but he was spoken to

firmly, sharply if necessary, but in a straight forward and essentially kindly fashion.

Some of the superior teachers were formal and precise, old-fashioned if you wish; while others tended to be more informal. But they were always in charge and in command of the situation. Pupils were given definite assignments, always knew what was expected of them, always knew where they stood. Pupils were held to high standards of achievement. It was clearly evident, in spite of popular opinion to the contrary, that children like better the teacher who is exacting, has definite goals and assignments, maintains high levels of accomplishment which are within the range of the pupils' capacities, provided that the teacher likes and respects them, and is strict not as a compensation for weakness but as a real indication of decisiveness and strength.

The objection might be raised that the teacher just described would be likely to dominate children and hence prevent them from becoming self-directing, independent and self-reliant. Such restrictions to development are not necessarily the result of assertiveness, strictness and formality on the part of the teacher, neither is it guaranteed that the passive, laissez faire teacher will help to cultivate independence and self-direction in children.

> Mr. BB. has the work in woodworking arranged in a series of models to be constructed, and when each object is completed he checks it over carefully for errors, and a mark is assigned on the number of errors which are found. This serves as a challenge to the boys which they meet with considerable enthusiasm. He holds them strictly to high standards, and they attempt to meet his expectations.

The superior teachers played fair with pupils so that the classes of the superior teacher showed their serenity by their calmness and freedom from restlessness and nervousness.

The superior teachers were secure in themselves so that they could afford to be casual and informal. They could dare to show a sense of humor, to enjoy a joke, to wisecrack a bit, to do a little clowning - and the pupils loved this informality.

> Miss U., a teacher of music, asked for another song. This time she stood in front of the class and joined the class in singing. She showed them what she wanted by clowning a bit with her voice and by the use of her facial expression. The children smiled at her as she burlesqued the piece slightly, then they joined in whole-heartedly.
>
> During one song about a ship crossing the ocean, she used her arms freely as part of the ritual of the song, and the class followed her lead, everyone singing, and gayly tossing their arms about.
>
> Then Miss U. passed out a number of instruments which made various sorts of noises, clicks, etc. used for keeping rhythm. The children vied for these, waving their hands frantically. Once these had been passed out, the class joined together in a Spanish-type song. The rhythm was perfect, the singing excellent and spirited, the whole performance was beautifully done.

The Inferior Teacher: Lacking Inner Strength

On the opposite side of the picture, the inferior teachers were characterized by expressions of their weakness which showed itself most clearly by a lack of sustained assertiveness. This group of teachers would consistently let small infractions of the rules and irregularities go by, apparently unnoticed, at least without protest, and then as the disorder and confusion rose they would lose their tempers and express their rising anger by shouting and screaming at the pupils, or meting out punishment.

There was emphasis on obedience and obeying the rules. Infractions were to be met by punishment. From the data there appeared a tendency for more of the inferior than of the superior teachers to make out detention slips for disorder in the corridors. As a result of the strain under which the inferior teachers labored, they tended to be tense, worried looking.

> Miss B. moves around the room looking over pupils' shoulders to see what they are doing. Carrying a harried look, she straightened out one pupil, smiled briefly to another, frowned at another. Later she took a stand at one side of the room and for several minutes surveyed the room with an air of authority. If a pupil was not absorbed in reading, she would stare at him intently as though to strike him with her glance. If the pupil should look up and see her staring at him, he would immediately get busy with something.

There is a common superstition that it is a sign of strength in a teacher to be punitive and a sign of weakness to avoid the use of punishment. There is also a common belief that effective teachers use more punishment than the less effective teachers. It was possible to secure data on this belief in connection with this study. Demerit slips were issued to students who are disorderly in passing to and from class by teachers who held stations in the corridors. These slips were filed and served as a basis, when a sufficient number have accumulated, for bringing a student to the atten-

tion of the student court. All teachers in the school except seven have this corridor duty so that it seemed fair to relate the measures of effectiveness which have been obtained on them with the number of slips which they signed over a stated period. Seven teachers were not included because their assignment was in the cafeteria so that the situation was not comparable to the teachers who had corridor duty.

The correlations obtained are given in Table XIII.

In spite of the unreliability of the various measures, the correlations are substantial. With 25 cases, a correlation of .395 is significant at the .05 level and of .502 is significant at the .01 level. The conclusion is inescapable that the better teachers use punishment less frequently.

Some of the inferior teachers apparently forestalled disorder by such practices as scolding, nagging, bullying, threatening, picking on individual students, the use of sarcasm, and ridicule. Each of these practices bespoke an underlying weakness and timidity.

> One boy whom Miss T., a mathematics teacher, is trying to persuade to finish his exercise started arguing with her, and she becomes snappish with him. She nags at the children to keep them at their work, and they are hardly respectful to her. One boy answers, "I know, lady."

There was an obvious feeling of inferiority and inadequacy with respect to their pupils. In some cases there was actual fear of the pupils. So a false front of boldness was adopted which took the forms mentioned above.

> Miss Z., a teacher of English, after a preliminary announcement which failed to arouse much interest, began talking about Macbeth. She spoke without notes, with a rare fluency, illustrating what she had to say with passages from the play, but her manner of speaking, however, was not condusive to easy listening for she talked very fast, hardly pausing between sentences.
>
> Shortly after she began to speak, a messenger brought in a note which she read aloud. "It's from Miss CC.," she said. "She wants all the children who helped her yesterday to report to her at 2:30. Now, who helped her yesterday?" Over half of the class raised their hands. Some pupils laughed, and someone yelled out, "liar."
>
> Miss Z. took little notice of this misbehavior and continued talking about Macbeth. The observer found it rather interesting despite her rapid speech, but the children found it otherwise. A child in the first row began to mimic her speech habits while her back was turned. The child behind him slapped at him; he whirled around and hit back. The child sitting near the observer was reading a magazine, and the child directly in front was manipulating a metal band. The girls were quiet, but apparently not all were attentive, for Miss Z. stopped speaking suddenly and stared at a girl in the fourth row. In a deliberate voice she said, "Do you know why I'm staring at you? Now put that away."
>
> A pupil in the front row raised his hand. Miss Z. looked right at him but refused to recognize him and continued with the story of Macbeth. The pupil waved his hand frantically for nearly twenty minutes without being recognized.
>
> A pupil began grappling with another boy which finally broke out into a full scale wrestling match. One boy picked the other off the ground and held him nearly over his shoulders. Miss Z. unleashed a scream at them and they separated.

> One has the feeling that Miss E., a physical education teacher, conducts her classes in a way calculated to make everyone aware of her superiority. Her bearing, manner and dress all betray this fact. She has the air of a Sargeant Major controlling a set of rookies. Some of the following are illustrations of this point.
>
> "Some of those sneakers could be a bit cleaner."
>
> In response to a question on how to hit a ball, Miss E. said sharply, "If you got rid of the gum perhaps you could try harder."
>
> Mr. J., a teacher of metal working, calls a boy who has a T-square in his hand to him and asks him in a severe threatening manner what a T-square is for. The boy attempts to explain, but Mr. J. is dissatisfied with the explanation and keeps pushing him for further answers. Finally it appears that the boy has been using the square to pry off the lift on the heel of his shoe. Mr. J. asks him if that is what the T-square is for, and then makes him say "No." The whole incident is designed to make the boy feel uncomfortable and ashamed.

In other cases the inner weakness showed itself directly by a passivity and lack of assertiveness. These teachers were ineffective because they could not exercise normal energy, vigor and assertiveness.

> Miss CC., an art teacher, was very forceful and her pupils were free, easy and informal. The observer believed that they might take advantage of her if things did not go right.
>
> When the next class came in, the pupils

TABLE XIII

COEFFICIENTS OF CORRELATION BETWEEN NUMBER OF DEMERIT SLIPS ISSUED AND MEASURES OF TEACHER EFFECTIVENESS

(A negative coefficient means that teacher effectiveness is negatively related to the number of demerit slips.)	
Principal Ratings	
Disciplinary Control	-.286
Relationships with Pupils	-.381
Success in Terms of Pupil Achievement	-.201
All Round Excellence	-.175
Average	-.381
Scale of Teacher Effectiveness	+.010
(A positive coefficient means that pupil estimate of a superior teacher is positively related to few demerit slips.)	
Pupil Rankings	
Question	
1	+.509
2	+.422
3	+.473
4	+.435
5	+.337
6	+.478
7	+.406
Average	+.467

were somewhat noisy and informal. One boy gave a shrill whistle. Miss CC. went around trying to get each pupil started on doing something. Some pupils got a little fresh with her. She asked one boy to collect the water saucers to which he objected, whereas another boy who wanted to do it was overlooked. While her back was turned, a boy snapped water from a faucet, another boy was singing rather obviously, two boys started wrestling while she was talking to a pupil about his painting.

Some of the inferior teachers gave the impression of being tired, monotonous, uninteresting, pedantic.

Miss B. reads aloud in a tired, monotonous voice a notice about a TV essay contest. She reads it as though it were something of great importance that every pupil should be interested in, although her own manner was anything but inspiring. Hardly a pupil seemed to be paying attention to her as she read.

The classes of the inferior teachers usually were pretty grim. The teacher took herself very seriously and dared not laugh or enjoy being with boys and girls. There was noticeable absence of a sense of humor. Or if they indulged in humor it had a sting to it, usually at the expense of some individual pupil.

Miss E. a teacher of dancing, stopped the class and with a humorous expression said, "Someone is not with us." She fixed her gaze on a girl in the back of the room. When most of the class was looking in the direction of the girl and tittering, Miss E. said, "Let's try it again, do you think you can keep with us." The girl blushed and nodded and the exercise continued. This sort of personal ridicule before the group is apparently a characteristic of Miss E. She used the technique several times in each class; and each time apparently enjoyed what she was doing.

These teachers took offense easily. Irregularities in the class were threats to their authority and prestige. They were very sensitive to possible meanness or ill-will among their pupils.

The Superior Teacher: Organization and Integration

With regard to the third point the evidence for better personality organization and integration among the superior teachers was not so evident. Although not revealed in the notes, there was an impression on reviewing the observation reports that the superior teachers organized their work somewhat more effectively. Goals of the instructor were better defined, and there was a more clear-cut time table for accomplishing these goals. The superior teachers were on the whole more businesslike.

Miss A., a teacher of social studies, is a solid, stable, businesslike, well-integrated person. One gets the impression that pupils do very little fooling in her class. After a film was shown she talked about it in a strong, firm voice, and pupils listened. There was absolute respect given to what she said. She was patient, did not move from a topic until the pupils had mastered it. She was a good organizer and had the best organized homeroom in the building. She apparently appeals to the more intelligent pupils in her classes and is not a warm person.

One gathers the impression that the superior teachers worked a little harder at their jobs and were really interested in it. Their work was not something to run away from at the end of a period or of a day, but they had invested more of themselves in their work.

The well-organized teacher can bend her endeavors toward helping children to become self-directing as well as the opposite. Teachers who organize their work effectively can help pupils to form purposes equally as well as to become dependent followers.

One of the superior teachers impressed the observer by her ability to make helpful explanations. Tasks were analyzed and laid out in one, two, three fashion, expressed in language that every boy or girl could understand. This ability to make clear explanations and to give clear directions betokened an orderly mind capable of analyzing a situation into its elements. And this teacher showed that this was not a routine performance for when she was expected to use a committee method in her class she gave her pupils clear and explicit directions as to how to organize work on a committee basis with the result that a high degree of skill in group activity was in evidence from the start.

The Inferior Teacher; Lack of Organization and Integration

The inferior teachers, on the other hand, were disorganized and lacked the evidences of personality integration. The goals of instruction were less clearly defined, and the methods for reaching goals were not so clearly outlined. The inferior teachers found it more difficult to keep to a time schedule. It was very clear that the pupils were more disorderly and their activities did not show the straight forward approach

to clearly defined goals.

> Miss P. is a problem teacher. Pupils say she doesn't know what she is doing. She is disorganized. When the team is all together, she is in left field. She gives the visual-aid people all sorts of trouble. She constantly overorders, and she never gets her orders in right. The school has a pretty rigid system about this, but she keeps changing her orders. Whenever the principal tells her about this, she just says, "Oh, I'm awfully sorry. I'll take it back and revise it." Then she does the same thing again.
>
> She is continually sending lists to the guidance department of failing students. Sometimes she changes the list several times within 24 hours.
>
> Pupils in her class were obviously bored. Many were visiting each other, there was always someone walking around the room, small conversations were going on everywhere, and a number of pupils spent time watching outside activities from the window. Two pupils were openly reading newspapers.

The inferior teachers were not so devoted to their work as the superior teachers. At the end of a period, they would heave a sigh of relief and show obvious pleasure at seeing the pupils depart. Preparing for a class was more of a chore.

> After the class was dismissed, Miss T. a mathematics teacher, gave a sigh of relief and visibly relaxed. She talked in a friendly way with some pupils who lingered behind.

The inferior teachers did not have the capacity to challenge pupils or to hold high standards of excellence. There was greater inconsistency in the inferior teachers. On one occasion there would be a show of zeal, pupils would be told what was expected of them. But if pupils failed to accomplish these requirements, they might be overlooked or there would be nagging, scolding or threatening.

Pupils gave definite signs of being ill at ease with the inferior teachers. They did not know precisely where they stood with them - and the uncertainty led to a dilettant, casual attitude toward study which led to restlessness and inability to concentrate.

> One did not feel that Miss E., teacher of physical education, was a part of the group but that she was a somewhat unapproachable leader. The girls did not cluster around her at the beginnings and ends of the periods. She seemed intent on order and lineups during times when the children normally can approach her. There was a good deal of tittering in her formation and the girls seemed restless when she tried to impose an order. She does not punish outright, and she is not inclined to punish every deviation. Thus the children feel it is permissible to do anything that they can get away with.

In concluding, a comparison of the results of this study with the factor analysis studies of Ryans[5] and Ryans and Wandt[6] will be made.

It should be noted that the Ryans-Wandt studies are only by implication studies of teacher effectiveness. What these investigators did, in short, was to have teachers rated on 18 characteristics that could be observed in the classroom (as well as securing the ratings of four items referring to pupil behavior). The intercorrelations of these ratings were computed, and a factor analysis of these correlations was made in order to determine the factors (relatively independent factors) that would account for these correlations. What this procedure does, in effect, is to boil the 18 characteristics down to three factors, taking into account the overlappings and duplications of the 18 traits. Jensen has translated the three factors which emerged from the factor analysis into behavior as "we see it in the classroom" in the following statement:

"....the teacher who demonstrates many of the behaviors contributing to Factor 1 is warm, friendly, and responsive in associations with pupils. She (or he) shows a liking and an affection for pupils without being unduly demonstrative. She appears approachable to all pupils, speaking to them as to an equal. In her relationships with students she exhibits tact, showing an awareness of a pupil's emotional problems and needs, and is alert to differences in individual abilities. She is tolerant of error on the part of pupils demonstrating patience and sympathy with a pupil's viewpoint. She resists any desire to dominate or to dictate a pupil's action and encourages pupils to make their decisions. She distributes her attention to many pupils and rotates leader-

5. D. G. Ryans. "A Study of Criterion Data of a Factor Analysis of Teacher Behavior in the Elementary School," Educational and Psychological Measurement, XII (1952), pp. 333-44.

6. D. G. Ryans and Edwin Wandt. "A Factor Analysis of Observed Teacher Behavior in the Secondary School," Educational and Psychological Measurement, XII (1952), pp. 574-586.

ship activities impartially. She finds good things to which to call attention in pupils. Her behavior with pupils is stable and predictable.

"The teacher who demonstrates many of the behaviors contributing to Factor 2 (and this teacher may be one who also shows behaviors contributing to Factors 1 and 3) is systematic in procedure and responsible in behavior. Her (or his) teaching shows evidence of careful planning, though the procedure is flexible enough to permit needed adaptions. She anticipates problems that might arise and is prepared to meet them. She demonstrates skill in holding a class discussion together, working toward desirable objectives. She provides for review and properly spaced learning. She is willing and eager to assume responsibilities in connection with her teaching, and is conscientious and punctual. She suggests aids to learning in providing "study hints." She controls unusual and difficult classroom situations. Her assignments are definite and she gives adequate directions. She calls attention to and emphasizes good work, insisting upon standards of quality. She is unruffled by problems developing in the classroom, and is dignified without being stiff or formal. She accepts criticism and is undisturbed and unembarrassed by it. She is realistic in approach, showing good common sense. She gives complete accurate, and satisfying answers to questions. She is constructively critical in approach to subject matter.

"Finally, the teacher who demonstrates behaviors contributing to Factor 3 challenges and stimulates pupils to do their best. She (or he) is clever and witty without being facetious. She is animated. She is well informed, showing evidence of imagination and independence in thought. She shows initiative in taking a new approach and resourcefulness in making an explanation or demonstration. She uses original, interesting, and sometimes relatively unique devices to aid instruction. She is flexible in adapting explanations and activities to pupil needs.

She takes advantage of pupils' questions to further clarify ideas. She brings lessons successfully to a climax, relating them to major objectives. She speaks of the future optimistically. She possesses a good sense of humor. She is constructively critical in approach to subject matter. She is unhurried in class activities, speaking quietly and slowly. She is alert to "pick-up" the class when pupils' attention shows signs of lagging. Her voice has an agreeable tone and interesting inflection. She is skillful and fluent in expression. She possesses personal charm, free from distracting personal habits. She is sure of herself and self-confident in relations with pupils. She has the classroom situations under control at all times.

"In summary, there is a need for a better understanding and definition of the patterns of personality traits and classroom behavior of teachers. As this behavior is observed in the classroom, three major areas emerge: (1) the teacher behavior that suggests warmth, friendliness, and responsiveness in associations with pupils, (2) the teacher behavior that appears to indicate systematic procedure and awareness of responsibility, and (3) the teacher behavior that seems to challenge and stimulate pupils to do their best[7].

It will be noted that the three factors which have come out of the study reported in the present paper (liking pupils, inner strength, personality organization) are inner factors having to do with feelings and personality organization. These inner factors exhibit themselves in characteristic behavior but actually these behaviors may vary from teacher to teacher. One teacher who possesses "inner strength" may show it by exercising firm control and not permitting irregularities to be passed by unnoticed. Another teacher, on the other hand, may show inner strength by considerable permissiveness, giving pupils freedom to express themselves in individual fashion. One teacher who lacks "inner strength" may attempt to control the situation by a show of authority and the excessive use of punishment; whereas another teacher lacking inner strength may be weak and passive, permitting the hostilities of pupils to gain the upper hand. Ryan's and Wandt's factors, on the other hand, are strictly observational factors.

It is the conviction of the writer that the effective (or ineffective) teacher cannot be defined in terms of behavior. Effectiveness can be expressed by many different kinds of behavior. It is believed that the basic determinants are to be found in the personality structure of the teacher rather than in outward behavior. It would be simple, of course, if effectiveness in teaching could be defined by outer behavior because one could then proceed to construct instruments to select persons who have the qualities that make for effectiveness in teaching. But the inner structure of personality does not lend itself to simple testing, and at the present time objective methods of determining the effective teacher have not been found. But the present study does demonstrate a method whereby the effective teacher can be located.

Conclusions

1. The study is two-fold. Part one describes a method for locating more effective as con-

7. A. C. Jensen. "Defining Effective Teacher Behavior," California Journal of Secondary Education, XXVII (1952), pp. 439-41. Actually this article is mistitled for there is nothing in the work of Ryans and Wandt which defines effectiveness in teaching.

trasted with less effective teachers by having pupils rank their teachers on seven bases.

2. Pupil rankings agree with each other with coefficients of correlation in the .70's, .80's, or low.90's. This indicates considerable halo effect in the rankings of the seven questions.
3. Pupil rankings correlate with principal ratings in the .60's for principal ratings of teacher discipline, in the .70's for teacher relationship with pupils and in the .40's for teacher ability to secure pupil achievement. These correlations are very satisfactory and shows good agreement when one takes into account the probable unreliability in the principal's ratings.
4. In the second part of the study based on the observation of teachers who were ranked high and those ranked low by pupils, the following three factors seemed to differentiate the teachers in the extreme groups:

a. Superior teachers liked children; inferior teachers disliked children.
b. Superior teachers were personally secure and self-assured; inferior teachers were personally insecure and had feelings of inferiority and inadequacy.
c. Superior teachers were well integrated and possessed good personality organization; while the inferior teachers tended to be personally disorganized.

SIGNIFICANT CHARACTERISTICS OF COLLEGE STUDENTS WITH VARYING FLEXIBILITY IN READING RATE: I. EYE-MOVEMENTS IN READING PROSE

FRANK LAYCOCK
Chico State College
Chico, California

Introduction

AN EARLY stage in learning to read is recognizing simple printed words and what they mean. With practice, recognition gradually embraces more complicated words and ideas. Then when some ideas can be understood more easily than others, a further skill is very helpful—matching speed to understanding, so that simple passages take less time to read than difficult ones. However, most people do not acquire this knack of flexibility; they maintain a constant rate regardless of its efficiency. Especially in college, where reading is the chief tool for gathering information, a student stuck in low gear is handicapped. Formulas, poems, short stories, and historical documents may all have to be understood, but understanding need not come at the same rate for all. Efficiency, therefore, implies knowing when to change speeds, and then being able to do so without undue trouble.

The Problem

In investigating voluntary change of pace, Anderson found that good readers shift readily from one reading situation to another; poor readers have trouble (1). For example, in going from details to general ideas, poor readers drop in speed and comprehension, while good ones maintain their efficiency. Poor readers seem to be most comfortable in their habitual reading pattern, and do not easily change. What would happen, however, if readers, all with similar average skill, were asked to shift speeds as quickly as they could? If they read simple enough passages, so that complexity or tension did not bother them, rate increases should vary with either visual skill or established habit. It is just this problem which is considered here: what happens to the oculo-motor behavior of good readers when they try to read simple material very quickly?

Previous Research

A few studies pertinent to flexibility in rate supplement Anderson's work and introduce this investigation. Years ago, Judd and Buswell noted that superficial and careful reading each produce different eye-movement patterns (2). Vernon claimed that reading is most irregular when disturbed by conflicting interests or emotional tension. Trying to read fast, for instance, may compete with the intention to fathom meaning, and thus produce irregularity (5:65). Tinker found that what the camera records as the fixation spot is only the approximate center of the visual field, and that when reading quickly a person can perceive the fringe about the center clearly enough to recognize words (3). If a reader is not accustomed to use the cues from this fringe, however, it may be hard for him to shift suddenly to them, for speed is partly governed by the attitude habitually associated with reading (4).

Subjects

The readers in this study were drawn from 391 applicants for admission to a cosmopolitan university following World War II. These sub-freshmen, and 101 upper division and graduate students then attending the same university, read several passages, making almost identical average rate and comprehension scores. The sub-freshmen also read at the same average rate similar material used with 680 applicants in previous years. They were predominantly men (84 percent) and averaged twenty-two years of age—both common post-war statistics. On standard aptitude tests like the Otis and the ACE their mean score was at the sixtieth percentile for the nation's college students. These sub-freshmen, then, were representative college students in their customary reading rate and tested academic aptitude.

Procedure and Apparatus

Two types of readers were selected: those who could read quickly when asked, and those who could not. Two equivalent reading selections were adapted from scripts used in a radio program which reports interesting facts gathered in university research. One selection dealt with little-known foods and minerals in the ocean,

the other with earthquake research. Each had about 2300 words and was checked for easy comprehension and high interest. Students were asked to read the first passage in order to answer simple questions afterwards ("normal" reading rate), and to read the second as fast as possible without missing important points which would later be checked ("advanced" rate). Generous warm-up time dissipated tension or confusion, and yielded excellent cooperation. A short quiz followed each selection, on the basis of which students with less than 75 percent comprehension were arbitrarily excluded. Each remaining reader's "normal" and "advanced" rates were computed, and the difference between the two rates taken as a measure of voluntary flexibility in reading rate.

Subjects were then arbitrarily divided into extreme groups by choosing for "flexible" readers 37 who were at least 1σ above the mean rate increase, and 35 "inflexible" readers 1σ below. Table I shows how these groups compare with each other, and whether the difference between them is significant, using the ratio Difference/$\sigma_{\text{Difference}}$.

These two types of readers, each satisfactory at normal rate but different in ability to read quickly, were studied for eye-movement characteristics, using standard techniques to derive calculations of rate, fixation frequency, fixation duration, and regressions. Three cards were prepared to use at the camera, each with a short text on it: a practice card to acquaint the reader with the camera situation, a card for "normal" reading (Card I), and one for "advanced" reading (Card II). Texts were comparable in length and difficulty, and questions afterwards tested comprehension. Reliability of the two cards for each of the measures calculated appears in Table II. Careful explanation introduced each reader to the aim of the study, and the practice card helped adjustment to reading at the camera. No subject afterwards reported any confusion or difficulty in following directions.

Results I: Differences Between Each Group in Reading Card I and Card II

What differences did the camera show between the eye-movements of flexible and inflexible readers? Table III summarizes the data for each of the measures, arranged according to the groups ("flexible" and "inflexible") and the cards read (I and II). Thus one may see how each group compared with the other in "normal" reading (Card I), and in "advanced" reading (Card II); and also how significant the differences are between the two groups.

Reading rate: In reading Card I, reading rates were faster than in the original screening (cf. Table I), probably because the passage in Card I is much shorter and somewhat easier to read (98 words, as compared with about 2300). There is a difference in mean rate between the two groups, perhaps because the "flexible" group automatically read a short, easy passage in such a way as to bring flexibility into play. In reading Card II, the same two facts appear: a general increase in rate over the original screening, and, as expected, a significant difference in rate between the two groups. Here one can see the striking difference in flexibility between the two groups.

Number of fixations: In reading Cards I and II, the flexible group was the more efficient, requiring significantly fewer fixations per passage. An indication of the difference in usable visual span comes in translating the mean number of fixations per group into letter-spaces: 70.5 and 78.7 fixations for Card I are equivalent to 7.2 and 6.5 letter-spaces respectively; 53.6 and 65.5 fixations for Card II, to 9.4 and 7.6 letter-spaces. Thus, cutting down the total number of fixations means increasing the amount of print taken in per fixation.

Duration of fixation: Reading rate and number of fixations combine to determine mean duration of fixation. The difference between the two groups shows the flexible reader made slightly longer fixations on Card I, but this difference is not significant. In shifting to Card II, this tendency disappeared. A plausible explanation would be that the need to grasp more material per fixation, as reflected in the widened fixation span, stimulated improved concentration and thereby facilitated speedier comprehension per fixation. Such an explanation shares Anderson's view that the poor reader is too engrossed in the mechanics of reading to concentrate on the content and thus permit faster comprehension.

Number of regressions: The flexible group tended to make fewer regressions in reading both cards, but the difference between groups was not dependable.

Discussion: The comparisons made above show that in each reading situation by itself, the flexible group was significantly superior to the inflexible in rate and fixation span. In fixation duration and regressions, the differences were not so reliable. These data bear out the contention that a group of satisfactory readers further segregated on the basis of flexibility differ very much as do readers originally segregated into satisfactory and unsatisfactory on the basis of general reading and academic skill.

Results II: Changes in Eye-Movement Characteristics from Card I to Card II

The major intent of this study was to probe

TABLE I

MEAN READING RATES IN SCREENING PASSAGES FOR TOTAL GROUP, FLEXIBLE GROUP, AND INFLEXIBLE GROUP
(Words per Minute)

Group	N	Mean Rate		Gain	Difference in Gains	σDiff.	Diff. / σDiff.	Significance
		Normal	Advanced					
Total	391	220.4	308.1	88.7				
Flexible	37	231.1	420.0	188.9	153.3	9.2	16.7	99+/100
Inflexible	35	219.3	255.0	35.7				

TABLE II

RELIABILITY COEFFICIENTS OF CAMERA CARDS I AND II

Measure	Card I			Card II		
	r	PE_r	N	r	PE_r	N
Total Fixation Time	0.92	0.02	71	0.97	0.01	71
Total Number Fixation	.83	.03	71	.94	.01	71
Mean Duration Fixation	.75	.05	71	.77	.04	71
Total Number Regression	.80	.03	71	.76	.04	71

TABLE III

COMPARISON OF FLEXIBLE AND INFLEXIBLE GROUPS IN EYE-MOVEMENT CHARACTERISTICS, CARDS I AND II

Measure	Group	N	Result		Diff.	σDiff.	Diff./σDiff.	Significance
Mean rate (words per minute)	Flexible	37	Card I	355.8	34.2	16.4	2.1	98/100
	Inflexible	35		321.6				
	Flexible	37	Card II	533.1	105.1	31.1	3.4	99+/100
	Inflexible	35		428.0				
Mean number fixations	Flexible	37	Card I	70.5	8.2	3.3	2.5	99/100
	Inflexible	35		78.7				
	Flexible	37	Card II	53.6	11.9	3.3	3.6	99+/100
	Inflexible	35		65.5				
Mean duration fixations (1/30 seconds)	Flexible	37	Card I	7.6	0.2	0.3	0.7	———
	Inflexible	35		7.4				
	Flexible	37	Card II	6.8	0.4	0.2	2.0	98/100
	Inflexible	35		7.2				
Mean number regressions	Flexible	37	Card I	9.5	1.4	1.8	0.9	———
	Inflexible	35		10.9				
	Flexible	37	Card II	6.4	0.8	1.2	0.7	———
	Inflexible	35		7.2				

TABLE IV

COMPARISON BETWEEN READING GROUPS OF CHANGES IN EYE-MOVEMENT CHARACTERISTICS FROM CARD I TO CARD II

Measure	Group	N	Card I	Card II	Change	Diff. in Change	σDiff.	$\frac{\text{Diff.}}{\sigma\text{Diff.}}$	Significance
Mean Rate (words per minute)	Flexible	37	355.8	532.1	177.3	69.7	19.3	3.6	99+/100
	Inflexible	35	321.6	428.0	106.4				
Mean Number Fixations	Flexible	37	70.5	53.6	16.9	3.7	2.3	1.6	95/100
	Inflexible	35	78.7	65.5	13.2				
Mean Duration Fixation (1/30 sec.)	Flexible	37	7.6	6.8	0.8	0.5	0.2	2.5	99/100
	Inflexible	35	7.4	7.1	0.3				
Mean Number Regressions	Flexible	37	9.5	6.4	3.1	0.6	1.1	0.5	———
	Inflexible	35	10.9	7.2	3.7				

the eye-movement characteristics of readers as they tried to read faster. After demonstrating the tendency toward superiority displayed by the flexible group in each reading situation separately, it is interesting to consider the way each measure changes as the reading purpose shifts. Table IV summarizes these changes.

Reading rate: The two groups differed most in ability to change rate, as would be expected, since that measure was the original criterion for segregation. Even when the "inflexible" group gained in rate, the flexible group gained more. The inflexible readers made a greater gain (nearly three times as much) than they had earlier in the long passages used for screening, probably because of the shorter and easier material in Cards I and II. The flexible group made a proportionately smaller gain, because they were closer to their speed ceiling and could not expect to triple the original rise.

Number of fixations: The change in number of fixations shown in Table IV is even more striking when translated into letter-spaces: the flexible group increased its average span by 2.2 letter-spaces, or 31 percent; the inflexible group by 1.1 spaces or 17 percent. These measures of span, of course, derive from reading connected discourse, not isolated tachistoscopic presentation, where possible recognition span is much greater. Over an extended period of reading, however, the margin of advantagle displayed by the flexible group is notable.

Duration of fixation: In mean duration of fixation a small numerical difference is important, too. Each group devoted less time to an individual fixation in order to develop speed, but the flexible group saved over twice as much time at each fixation. The flexible group's initial disadvantage on Card I disappeared on Card II.

Number of regressions: The least reliable difference between groups occurred in number of regressions. The inflexible group showed greater improvement, but still averaged more regressions on both Cards I and II.

Results III: Comparison of Reading Groups Separately for Changes from Card I to Card II

Before summarizing the changes made in shifting reading purpose, a limitation in the investigation needs comment. It has been necessary to point out that the inflexible readers made gains, even though not to the extent shown by the flexible readers. In Table V the amount and significance of the gains each group made by itself appear to supplement the foregoing comparisons between groups.

Reading rate: In going from Card I to Card II, the inflexible group as well as the flexible group made a gain which is thoroughly significant. To that extent Cards I and II apparently do not duplicate the original screeing comparisons, where the inflexible groups made a much smaller average gain. The difference between the screening passages and the camera card results probably arose because the camera cards were shorter (100 words rather than 2300) and easier (mean rate well over 100 words per minute faster), and because a briefer reading time allowed a greater spurt.

Number of fixations: Here the pattern of significant gains in each group's scores continued, with the flexible group still relatively superior.

Duration of fixation and number of regressions: For each of these measures there was the same tendency for each group to improve, but the flexible group's degree of superiority was not so great as in rate and number of fixations.

Comment: The fact that the inflexible readers made greater improvement at the camera than in the original screening suggests that using the words "flexible" and "inflexible" to describe the groups is faulty. They are more or less flexible, rather than flexible or inflexible. The greatest and most consistent differences between the groups lie in rate and number of fixations; less clear are the differences in duration of fixation, and insignificant those for regressions. Table VI shows in detail how significantly each group changed in the measures computed. Table VII indicates the degree to which the flexible group made superior scores on each of the cards and in shifting from one card to the other. This superiority is stable for rate and number of fixations.

Conclusions

This investigation suggests certain conclusions.

1. In a group of readers whose comprehension and average reading rate are good, some will be more flexible in reading rate than others, and the two types can be segregated by merely asking them to read faster. In this study, each group thus segregated showed some flexibility with simple passages, but the more flexible group was significantly superior in most eye-movement measures. In reading longer and more difficult passages the divergence in flexibility increased. Hence on fairly difficult material, like that in college work, greater flexibility is a desirable goal for many readers who already comprehend well, and whose average is good.

2. Rate of reading, however, is not the all-in-all, for fast reading still demands comprehension. The more flexible reader is merely one who can jump from one reading situation to another and read in each at a faster or slower

TABLE V

COMPARISON FOR READING GROUPS SEPARATELY OF CHANGES IN EYE-MOVEMENT CHARACTERISTICS FROM CARD I TO CARD II

Measure	Group	N	Card	Result	Diff.	σDiff.	$\frac{\text{Diff.}}{\sigma\text{Diff.}}$	Significance
Mean rate (words per minute)	Flexible	37	I	355.8	177.3	25.5	7.0	99+/100
			II	533.1				
	Inflexible	35	I	321.6	106.4	24.2	4.4	99+/100
			II	428.0				
Mean number fixations	Flexible	37	I	70.5	16.9	3.1	5.5	99+/100
			II	53.6				
	Inflexible	35	I	78.7	13.2	3.4	3.9	99+/100
			II	65.5				
Mean duration fixations (1/30 seconds)	Flexible	37	I	7.6	0.8	0.2	3.8	99+/100
			II	6.8				
	Inflexible	35	I	7.4	0.3	0.2	1.5	93/100
			II	7.1				
Mean number regressions	Flexible	37	I	9.5	3.1	1.2	2.6	99+/100
			II	6.4				
	Inflexible	35	I	10.9	3.7	1.6	2.3	99/100
			II	7.2				

TABLE VI

COMPARISON OF EXTENT OF DIFFERENCE BETWEEN CARDS I AND II FOR FLEXIBLE AND INFLEXIBLE READERS

	$\frac{\text{Diff.}_{\text{I-II}}}{\sigma\text{Diff.}_{\text{I-II}}}$			
Group	Rate	Number Fixations	Duration Fixations	Number Regressions
Flexible	7.0	5.5	3.8	2.6
Inflexible	4.4	3.9	1.3	2.3

TABLE VII

COMPARISON OF EXTENT OF SUPERIORITY OF FLEXIBLE OVER INFLEXIBLE GROUPS FOR EYE-MOVEMENT MEASURES IN READING CARDS I AND II

	$\frac{\text{Diff.}_{\text{(Flex.-Inflex.)}}}{\sigma\text{Diff.}_{\text{(Flex.-Inflex.)}}}$			
Comparison	Rate	Number Fixations	Duration Fixations	Number Regressions
Card I	2.1	2.5	0.8*	0.9
Card II	3.4	3.6	2.0	0.7
Change: I to II	3.6	3.7	2.5	0.7

*Inflexible greater than flexible

rate than usual, depending on comprehension and significance. The less flexible reader, on the other hand, does all his reading—regardless of speed or ease of comprehension—at more nearly the same rate. He does not comfortably widen the span of a single fixation or skip unimportant words. If comprehension is adequate, as it was in this study, it is possible that the difference between more and less flexible readers reflects the degree of skill in controlling eye-movements. Less flexible readers commonly report that they cannot successfully think of increasing visual span, for example, while still maintaining adequate comprehension.

3. Reading improvement programs at higher levels like the college may well segregate students on the basis of flexibility for at least part of their work, so that the more proficient do not plod through the same training as their less versatile fellows. For the less flexible reader who comprehends well, attention to eye-movements themselves may be profitable. Tachistoscopic training in quick recognition, for instance, may help widen fixation span and improve reading accuracy and flexibility.

4. Since the most significant differences found between the more and less flexible readers were rate and fixation span, these eye-movement characteristics should be measured in a situation where comprehension is negligible. The extent to which flexibility depends on eye-movements would then show up more clearly.

Summary

This investigation segregated from 391 college students 37 who could shift reading rate easily and 35 who could not. At an eye-movement camera each group read a short passage (Card I) at normal rate, and a second (Card II) as fast as possible without lowering comprehension. Film records were analyzed for rate, fixation span, duration of fixation, and regressions.

When records for each card were examined separately, the flexible group's eye-movement characteristics were nearly all significantly superior, especially rate and fixation span. In measuring the changes in characteristics each group made in shifting from one card to the next, the same superiority appeared. A final analysis showed that the inflexible group improved its reading somewhat in going from Card I to Card II, although not so much as the flexible group did. As a consequence, the terms "more flexible" and "less flexible" are more appropriate.

Since comprehension for each group was adequate, one implication is that ease in controlling eye-movements is a factor in achieving flexibility. At higher reading levels, as in college work, it is wise to improve flexibility of rate not only by building a background for comprehension but also by developing skill in eye-movement control. This study suggests particularly that rate of progression and width of fixation are the eye-movement characteristics which training programs should emphasize.

REFERENCES

1. Anderson, Irving H. "Studies in the Eye-Movements of Good and Poor Readers," Psychological Monographs, XLVIII, No. 5 (Princeton: Psychological Review Co., 1937).
2. Judd, Charles H., and Buswell, Guy T. "Silent Reading: A Study of the Various Types," Supplementary Education Monographs, No. 23 (Chicago: University of Chicago Press, 1922).
3. Tinker, Miles A. "The Use and Limitations of Eye-Movement Measures in Reading," Psychological Review, XL (1933), p. 382.
4. Traxler, Arthur E. Ten Years of Research in Reading (New York: Educational Records Bureau, 1941).
5. Vernon, Magdelen D. The Experimental Study of Reading (Cambridge: Cambridge University Press, 1931).

II. MOTOR AND PERCEPTUAL SKILL IN "READING" MATERIAL WHOSE MEANING IS UNIMPORTANT

FRANK LAYCOCK

Introduction

PART I of this report described a group of college students who all had similar general reading ability. Within this group were some who could increase speed easily when reading simple prose ("flexible" readers) and others who could not ("inflexible" readers). The flexible readers differed significantly from the inflexible in their characteristic eye-movement patterns, particularly in fixation span and rate. Since the prose was very easy to understand, it should have been simple for all these students to read it quickly. Therefore the wide differences which appeared were probably not related to factors which make for easy understanding: familiarity, vocabulary, general background. Further, all the students had recently been tested by refractionists who found their apparatus satisfactory. Finally, there was no indication that confusion, tension, or unfamiliarity affected the reading measures. In order to find out what was responsible for the wide variation in rate increase, the same students were put in quasi-reading situations which eliminated or minimized concern for meaning, consecutive thought, or judgment. They would then use their eyes as they usually do in reading, and perceptual-motor habits would persist without any complicating demand for understanding meaning. Would the two groups of readers still show significantly different behavior?

The Problem

Three kinds of visual activity were chosen for measurement, to see if there was any difference between the more and less flexible readers. They were: (1) fixating accurately on a series of numbers placed irregularly along several lines of print to simulate a paragraph, with no need for comprehension or memory to blur concentration on accurate placement; (2) looking at successibely wider prose stimuli to see how much width could be "taken in" during very short exposures; and (3) checking pairs of unfamiliar but uncomplicated names to note differences in spelling or punctuation. Any differences between the two types of readers might give further help in explaining why one group could increase its reading rate so much more easily than the other.

Previous Research

Nothing in the literature considers the problem of flexibility in reading rate as it is related to motor or perceptual skill, but there are some obliquely related studies. Arnold and Tinker (2) draw a distinction between the short time in which a letter can be seen when shown tachistoscopically and the longer time devoted to a single fixation in ordinary reading. They report that when a fixation in reading lasts longer than 0.1 second there are probably other factors at work beside the requirements of clear perception: reaction time, apprehension and assimilation, time for pausing and moving on, speed of binocular coordination after the inter-fixation movement. Vernon (6) used a tachistoscope to present several short words in disconnected groups, in sentences and in connected prose. She found that errors are either typographical (especially in disconnected words) or meaningful (associated with changes in interpretation according to the meaning of the words). When meaning is clear there are few errors in reporting what has been seen; when not clear, the reports are inaccurate. Gates (3) found that efficient perception is reading is made up of special abilities rather than a single, general perceptual ability; therefore measuring perceptual ability in reading is better done with tests that use words than with tests employing designs or pictures. Tinker (5) used eye-movement photography to show that the center of the visual field for any given fixation is not the only spot where vision is clear enough to distinguish words. The less clear fringe is also useful for reading.

According to these scattered references, perceptual skill in reading should be measured in situations which involve or duplicate reading. Further, in the ordinary reading act it is hard to separate sensation from perception, particularly where grasping meaning is part of the process. The question, then, is whether activities which minimize comprehension without destroying the reading-like atmosphere can give any clues about motor or perceptual skill in reading.

Procedure and Apparatus

The flexible and inflexible readers described in Part I were used as subjects, to see if their

eye-movement characteristics when reading prose corresponded to differences in motor and perceptual skills which demand little comprehension.

Eye-movement records: At the eye-movement camera, each subject was asked to "read" two cards, each with 25 numbers scattered over five horizontal lines. Each card had single digits in erial progression (1, 2, 3, ...) from 1 to 9 and repeated, spaced at irregular intervals along the five lines. Directions emphasized accurate gaze at each number in turn, and made no mention of speed.

Tachistoscopic records: Three film strips were developed for projection through a specially designed tachistoscope onto a screen. The tachistoscope consists of two film strip projectors synchronized so that the image from one can be superimposed upon that of the other. Venetian-blind-like shutters in front of each projector lens work together, so that when the light from one projector shows on the screen the light from the other is automatically cut off. Thus one projector can show a constant field of light on the screen; the second projector can show its image for from 0.01 to 0.1 second, and the original field of light can then return—all so smoothly that to the eye no change in background illumination is apparent. In this investigation the constant background included a small dot at the center, to act as a guide for awaiting what would appear momentarily on the screen. This combination of constant background and reference dot is a tachistoscopic refinement which eliminates any difficulty in focusing the eye, adapting to illumination change (as when a bright light suddenly appears on a dark surface), or knowing exactly where to look. The only complication this double-tachistoscope imposes is preparing parallel film strips with equal brightness for each pair of frames.

Three different series of film strips were photographed. The first was a series of thirteen letters, so arranged that each letter was distant from the center by one more letter space than its predecessor; i.e., the first pair of letters had one vacant space between, the second pair three spaces, and so on, to a maximum of two letters separated by 25 spaces, or the equivalent in length of a 27-letter word. Each pair of letters had a clearly discernible small dot placed half-way between. The subject was asked to look steadily at the dot (which appeared uninterruptedly on the screen), and to identify whatever letters appeared by looking out of the "corners" of his eyes. The second series consisted of gradually lengthened words, phrases, and clauses: first a single letter, then a two-letter word, and so on through 25 steps to a five-word clause which contained 25 letter spaces. The third series was a combination of the ideas in the first two: sentences divided into two equal parts, one placed to the left of the center dot, the other to the right. This series started with single letters, one on each side of the dot; then a pair of two-letter words, one word on each side; and so on for 25 split sentences. All the sentences were sensible, easily understood, and devoid of long or awkward words. For the first two series the students were told to look at the screen and write down what appeared; the exposure time was 0.1 second. In the third series, they looked at each half of a sentence in order; total exposure time per frame was 0.4 second. Exposure times were determined empirically in a pilot study.

Name-checking accuracy: The last instrument was adapted from an Australian test of clerical accuracy. It consists of 50 paired names to be checked for similarity in appearance. Some of the pairs are identical in spelling, punctuation, and format; others have slight differences; and each item is comparable to the others in difficulty. The intent, as in many such clerical tests, is to check each pair of names for differences, and to finish as quickly as possible. The names, taken from Melbourne telephone directories, are those of firms and people not likely to be familiar to American students. Such a set of names demands the ability to detect slight differences in appearance, without either departing from a reading-like situation or giving undue advantage to readers already accustomed to the names or sequences.

Table I shows the reliability coefficients of the several measures taken. They are satisfactory for group comparisons, although the name-checking test results need to be used with caution.

Results I: Cards A and B

The number cards A and B were not used to measure flexibility in reading rate, but to see the degree to which each student could control his eye-movements when the demands of comprehension were small. Table II shows the eye-movement characteristics of the flexible and inflexible groups for these two cards.

Rate: The flexible group read the cards more quickly (note that the speed, expressed in 1/30 second units, increases as the number of units decreases). This superiority in speed may be due to a greater degree of control over eye-movements, since each group concentrated not on speed but on accurate fixations.

Number of fixations: Although each group averaged more than one fixation for the 25 digits per card, the flexible group was much more accurate. Making a few extra fixations is understandable, considering the fine movements

TABLE I

RELIABILITY COEFFICIENTS OF CARDS A AND B; TACHISTOSCOPIC FILMS 1, 2, AND 3; AND NAME-CHECKING TEST

Measure		r	PE_r	N
Cards A - B	Total Fixation Time	0.97	0.004	71
	Total No. Fixations	.86	.02	71
	Mean Duration Fixations	.86	.02	71
	Total No. Regressions	.79	.03	71
Tachistoscopic Films	No. 1	.82	.03	72
	No. 2	.77	.03	72
	No. 3	.85	.02	72
Name Checking List	Total Time	.87	.20	72
	Total No. Correct	.72	.40	72

TABLE II

COMPARISON OF FLEXIBLE AND INFLEXIBLE READING GROUPS IN EYE-MOVEMENT CHARACTERISTICS, CARDS A AND B

Measure	Group	N	Result		Diff.	σDiff.	$\frac{\text{Diff.}}{\sigma\text{Diff.}}$	Significance
Mean rate (1/30 sec.)	Flexible	36	Card A	233.9	33.0	14.4	2.3	99/100
	Inflexible	35		266.9				
	Flexible	36	Card B	218.3	25.2	14.3	1.8	96/100
	Inflexible	35		243.5				
Mean number fixations	Flexible	36	Card A	27.1	1.9	0.8	2.4	99/100
	Inflexible	35		29.0				
	Flexible	36	Card B	26.8	1.8	0.8	2.3	99/100
	Inflexible	35		28.6				
Mean duration fixation (1/30 sec.)	Flexible	36	Card A	8.2	1.0	0.5	2.0	98/100
	Inflexible	35		9.2				
	Flexible	36	Card B	8.2	1.0	0.5	2.0	98/100
	Inflexible	35		9.2				
Mean number regressions	Flexible	36	Card A	2.3	1.3	1.1	1.2	———
	Inflexible	35		3.6				
	Flexible	36	Card B	1.9	0.6	0.5	1.2	———
	Inflexible	35		2.5				

TABLE III

COMPARISON OF FLEXIBLE AND INFLEXIBLE GROUPS FOR SEVERAL MEASURES, TACHISTOSCOPIC FILM NO. 1 (EXPOSURE TIME PER FRAME 0.1 SECOND)

Mean Measure	Group	N	Result	Diff.	σDiff.	$\frac{\text{Diff.}}{\sigma\text{Diff.}}$	Significance
No. items correct	Flexible	37	9.1	0.9	0.5	1.8	96/100
	Inflexible	33	8.2				
No. letter-spaces preceding first error	Flexible	37	6.7	1.1	1.1	1.0	———
	Inflexible	33	5.6				
No. letter-spaces in better of best two consecutively correct	Flexible	37	12.3	0.8	0.7	1.1	———
	Inflexible	33	11.5				
No. letter-spaces in median of best three consecutively correct	Flexible	37	10.4	0.7	1.0	0.7	———
	Inflexible	33	9.7				

necessary to change a fixation.

Duration of fixations: The flexible group spent consistently less time on an individual fixation, although there was neither specific instruction to go quickly nor any meaning to comprehend. This tendency suggests again superior control of the eyes.

Regressions: Although not significant, there is a tendency for the flexible group to need fewer approaches to a goal for accurate placement. Since there is none of the trouble over meaning which causes many regressions in ordinary reading, the regression here may be interpreted as related to the total number of fixations, which yielded a highly significant comparison.

Discussion: The pattern of differences between the flexible and inflexible groups is similar to that found in analyzing prose records in Part I of this study. Both reading situations show that rate and number of fixations (which correspond to fixation span in prose reading) are significant measures. When meaning is absent, fixation duration is more significant. Regressions tend to differentiate the two groups, but within each group variability is too high for reliability. Since the flexibility readers make a better showing in both prose and number reading, it is possible that the motor aspects of reading are important determiners of change in pace. This possibility supports Anderson's contention that poor readers are too engrossed in the mechanical part of reading to allow increases in speed (1:17). Such a conclusion is all the more compelling, because the inflexible readers were not poor readers, but just as effective in normal reading as their flexible companions. Apparently, good normal reading skill does not automatically bring flexibility in rate, any more than poor reading does.

Results II: Tachistoscopic Records

From the eye-movement camera, attention now turns to the way in which prose elements may be grasped in short exposures. Here comparisons are made between the two groups already segregated on the basis of flexibility in rate in reading prose, to see if superiority in one situation is related to performance in another.

Film no. 1: The first film presented pairs of letters which gradually spread farther apart to a maximum width of 25 letter spaces. This film drew on the ability to see letters in the periphery of vision, since students focused on the dot placed midway between each pair of letters. Table III gives the data derived from this film. The first measure, number of items correct, shows the flexible group to be superior at quick recognition. Skill like this is useful in skimming over a page of print. The accuracy score, however, treats all fifteen items in Film no. 1 as comparable units, whereas each item was wider than its predecessor. Saying that the flexible group, for instance, interpreted correctly an average of 9.1 units is not the same as saying that the maximum width seen correctly was that of the ninth pair of letters exposed. Few of the readers escaped making scattered errors before reaching a point where consistent errors began. The remaining entries in Table III are various estimates of the point where errors began, and yield quite different scores. The second and third scores are logically the best, since a random error followed by two or three correct exposures would arbitrarily assign a much lower score with the first estimate, and give a spuriously low estimate of ability. None of these three measures discriminated significantly between the two groups of readers, although there was a consistent tendency in favor of the flexible readers.

Film no. 2: The second film introduced words instead of isolated letters, thus giving subjects a chance to identify by means of configuration as well as individual letters. Table IV shows the accuracy scores for the two groups of readers. Although not quite significant, there is a strong tendency for the flexible group to see more in a fixation. Only one measure of width perceived was necessary, because seeing several words accurately is less likely to be the result of pure guessing than trying to pick out separate letters.

Film no. 3: In this film each student looked first at one side of a sentence and then at the other, all in 0.4 second. Table V implies that the lengthened exposure time did not bring a corresponding rise in score. The difficulty in shifting the eyes from one spot to another more than used up the additional exposure time. It is notable that the most reliable differences between groups appear in the third film, which is most like normal reading. In the others, differences occur consistently, although not to the same extent.

Tachistoscopic vs. ordinary perception: Table VI is a comparison of three essentially similar types of reading material: the prose camera cards (no. 1, normal reading rate; no. 2, accelerated reading rate) in Part I of this study; and tachistoscopic film no. 2. These situations resemble three positions along a continuum of increasing urgency in reading. Table VI clearly shows how much more can be seen in a fixation when presented tachistoscopically than even conscious acceleration permits in reading. Table VII summarizes a comparison between the data for flexible-inflexible readers' scores combined and parallel data from Robinson's investigation (4). These two tables indicate how easy it is to perceive written material in a short ex-

TABLE IV

COMPARISON OF FLEXIBLE AND INFLEXIBLE GROUPS FOR ACCURACY, TACHISTOSCOPIC FILM NO. 2 (EXPOSURE TIME 0.1 SECOND)

Mean Measure	Group	N	Result	Diff.	σDiff.	$\frac{\text{Diff.}}{\sigma\text{Diff.}}$	Significance
No items correct	Flexible	37	14.2	1.1	0.8	1.4	——
	Inflexible	33	13.1				
No. letter-spaces in widest sentence correctly seen	Flexible	37	18.5	1.3	0.9	1.4	——
	Inflexible	33	17.2				

TABLE V

COMPARISON OF FLEXIBLE AND INFLEXIBLE GROUPS FOR ACCURACY, TACHISTOSCOPIC FILM NO. 3 (EXPOSURE TIME 0.4 SECOND)

Mean Measure	Group	N	Result	Diff.	σDiff.	$\frac{\text{Diff.}}{\sigma\text{Diff.}}$	Significance
No. items correct	Flexible	37	4.8	1.1	0.6	1.8	97/100
	Inflexible	33	3.7				
No. letter-spaces in each half of widest sentence accurately reported	Flexible	37	7.8	1.8	1.1	1.6	95/100
	Inflexible	33	6.0				

TABLE VI

MEAN NUMBER OF LETTER-SPACES IN PROSE CARDS I AND II AND TACHISTOSCOPIC FILM NO. 2, FOR FLEXIBLE AND INFLEXIBLE GROUPS, EQUATED TO 0.1 SECOND EXPOSURE TIME

	Mean No. Letter-Spaces				
Group	Tachistoscopic Film No. 2	Prose Card I	% I/2	Prose Card II	% II/2
Flexible	18.5	2.2	12.0	4.5	24.3
Inflexible	17.2	1.8	10.5	3.1	18.0

TABLE VII

MEAN NUMBER WORDS PER FIXATION SEEN BY COMBINED FLEXIBLE AND INFLEXIBLE GROUPS IN PROSE CARDS I AND II AND TACHISTOSCOPIC FILM NO. 2, COMPARED WITH ROBINSON'S DATA

	Mean No. Words per Fixation				
Group	Prose Card I	Prose Card II	Tachistoscopic Film No. 3	Words per Fixation	Words per Tach. Exp. (0.1 sec.)
Flexible + Inflexible	1.3	1.7	4.6		
Robinson				1.6	4.0

TABLE VIII

COMPARISON FOR FLEXIBLE AND INFLEXIBLE GROUPS OF NUMBER OF LETTER-SPACES SEEN IN TACHISTOSCOPIC FILMS NO. 2 AND NO. 3, EQUATED TO 0.1 SECOND EXPOSURE TIME

Group	Mean Number Words Perceived 0.1 Second		
	Film No. 2	Film No. 3	% 3/2
Flexible	18.5	3.9	21.0
Inflexible	17.2	3.0	17.4
Flexible + Inflexible	17.8	3.4	19.1

TABLE IX

COMPARISON OF FLEXIBLE AND INFLEXIBLE GROUPS FOR MEAN NUMBER OF ITEMS CORRECT IN NAME-CHECKING TEST

Group	N	Mean No. Items Cor.	Diff.	σDiff.	$\frac{\text{Diff.}}{\sigma\text{Diff.}}$	Significance
Flexible	36	44.4	1.4	0.8	1.8	96/100
Inflexible	35	43.0				

posure when attention can be given over completely to one stimulus; and how much less can be grasped when there are the added requirements of smooth visual progression, accurate placement, and insertion of a new perception into the whole evolving pattern. The part played by shifting from one fixation to another appears in Table VIII. Once the inter-fixation movement complicates perception, the amount perceived is affected, for flexible and inflexible readers alike.

Results III: Name-Checking Test

The name-checking test required attention to small details of spelling and punctuation with no worry over meaning. Table IX shows that the flexible group was significantly superior to the inflexible, even though mean scores were not far apart. In the test, each student worked as fast as he could, and time taken was recorded. But there was no difference in speed between groups beyond that expected from chance.

Conclusions

College students who read satisfactorily under normal conditions may be segregated into two extreme groups on the basis of their ability to make themselves suddenly read fast. These two groups not only differ markedly in the way they read meaningful prose (cf., Part I of this study), but also display divergent characteristics when they "read" material which has little or no meaning. The more flexible reader, as compared with the less flexible, deals with this low-meaning material in such a way as to justify the following conclusions:

1. In looking successively at digits placed irregularly to form a "paragraph", he takes less time (even though not trying for speed), fixates each digit more accurately, requires fewer fixations to complete a set of digits, and tends to make fewer regressions. These are the same characteristics he displays when reading prose for comprehension.
2. The flexible reader can identify letters more accurately when given short glimpses of them away from the center of the visual field. It is this ability to use stimuli from the fringe of vision which enables him to widen his fixation span in prose reading and thus work more efficiently.
3. He is more accurate in catching sentences flashed momentarily on a screen, both when one fixation alone is allowed and when it is possible to make a rapid jump from one part of a sentence to another.
4. The flexible reader is more successful in identifying slight changes in spelling and punctuation, although not necessarily doing so more quickly.

From the broad view, these findings show that reading flexibility is associated with factors which can be detected by eye-movement camera, tachistoscope, and measures of discrimination. Rate and number of fixations are the two eye-movement characteristics most significant at the camera; accuracy is the measure most directly made by tachistoscope and discrimination test. The fact that readers show the same characteristics whether they are dealing with meaning or not, strongly suggests that a significant factor in reading flexibility is habitual patterns of motor and perceptual behavior. Efforts to direct an inflexible reader's attention to these habitual practices, and specific help in improving them are logical recommendations for training programs.

Summary

The 37 flexible and 35 inflexible college readers segregated in Part I of this study "read" three types of material which minimized or eliminated the need to gather meaning: (1) At the eye-movement camera they looked at two cards, each with 25 single digits spaced irregularly to simulate a paragraph, and tried to glance squarely at each digit in turn. (2) On a screen they observed tachistoscopically exposed print of three sorts: pairs of letters exposed farther and farther apart which the student tried to identify while gazing fixedly at a dot midway between; gradually lengthened sentences; and similar sentences exposed so that one half appeared at the left and the other half simultaneously at the right. (3) They marked 50 pairs of names to indicate slight differences between members of a pair in spelling or punctuation.

In these three situations, no mention was made of varying rate of reading. Yet in nearly all measures the group already designated flexible was significantly superior to the other. The flexible readers more quickly and accurately fixated digits at the camera; identified more letters seen in the "corners" of the eyes; caught more sentences flashed on the screen; and noted more small discrepancies in format. It is important to bear in mind that both groups of readers had similar, satisfactory mean reading rates in ordinary reading situations.

Since there was little meaning in this "reading," all the students' eyes had been checked and found satisfactory, and nobody reported confusion or tension; it appears that habitual motor and perceptual behavior is an important factor in reading flexibility. Students who increase speed easily seem to have eye-movement and perceptual habits which aid them; other students

seem to be bound to a more sluggish pace. Training programs should be able to aid less flexible readers by paying specific attention to these motor and perceptual habits.

REFERENCES

1. Anderson, Irving H. "Studies in the Eye-Movements of Good and Poor Readers," Psychological Monographs, XLVIII, No. 5 (Princeton: Psychological Review Co., 1937).
2. Arnold, Donald C., and Tinker, Miles A. "The Fixational Pause of the Eyes," Journal of Experimental Psychology, XXV (1939), pp. 271-80.
3. Gates, Arthur I. "A Study of the Role of Visual Perception, Intelligence, and Certain Associative Processes in Reading and Spelling," Journal of Educational Psychology, XVII (1926), pp. 433-45.
4. Robinson, Francis P. "The Tachistoscope as a Measure of Reading Perception," American Journal of Psychology, XLVI (1934), pp. 132-35.
5. Tinker, Miles A. "The Use and Limitations of Eye-Movement Measures in Reading," Psychological Review, XL (1933), pp. 381-87.
6. Vernon, Magdelen D. "The Errors Made in Reading," in Medical Council Report, No. 130 (London: British Medical Council, 1929).

AN EVALUATION INSTRUMENT FOR THE ANALYSIS OF TEACHER EFFECTIVENESS

GRANVILLE B. JOHNSON, Jr.
Emory University
Emory University, Georgia

Introduction

THAT THE prediction of teaching efficiency and evaluation of teaching success is of decided current interest is evidenced by the number of recent investigations related to the area (3).

Although considerable attention has been given the area and much excellent work has been done, the ability to evaluate efficiency and to predict teacher ability still has a rather wide margin of error.

The present study which is the result of four years of experiment does not presume to be the solution to the problem, but was developed with an approach which may be of interest to other investigators.

Problem

The problem was to ascertain the relationship between qualities possessed by teachers and teacher effectiveness, with the aim of predicting teacher success.

The steps of this problem may be subsumed in the following:

I. To devise an evaluation instrument for the analysis of teacher effectiveness in an actual teaching situation.
II. To utilize instruments which measure, at a molar level, the attitudes particularly relevant to teaching effectiveness.
III. To ascertain the relationship between teacher personality as revealed by the selected measuring instruments, and teacher effectiveness as revealed by overt behavior with the aim of predicting teacher success.

Problem I: The first step in the present investigation was to devise an evaluation instrument for the analysis of teacher effectiveness in an actual teaching situation.

A review of the literature reveals three primary approaches to measuring teaching effectiveness:

1. Evaluation of qualities assumed to function in the teaching act. (The relationship between teacher efficiency and specific teacher qualities has yet to be determined.)

2. Appraisal of teaching activity. (As the above, it appears logical that certain activities should be related to teaching efficiency.)

3. Evaluation of pupil growth. (It is generally concluded that this is the ultimate criterion of teacher effectiveness since it is a measure of what the teacher has done. Devising tools for testing the total desired change, however, precludes success at present (2).)

It may be observed from the analysis of methods of approaching the problem that teacher effectiveness may be defined in as many ways as there are ways of adjudging it. The definition of teacher effectiveness in the present study may be found in the following pre-suppositions and hypotheses basic to the justification of the proposed research. It is hypothesized that there is a positive relationship between teacher effectiveness (independent variable) and pupil growth (dependent variable); it is presumed that the greater the psychological insight the teacher has into principles of individual differences, human growth and development, and learning, and the greater the transfer of these principles she makes in practice, the greater will be the resulting growth. It is understood that different situations would demand different teaching techniques, i.e., there is no set technique which would be best fitted to every learning situation—cognizance of and adjustment to individual and group differences (intervening variables) would be a criterion of teaching effectiveness.

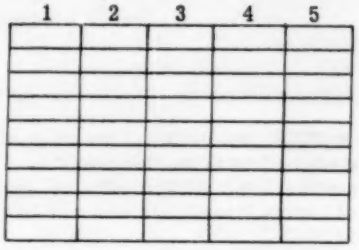
The relationship between the independent and the dependent variables would be 1.00 if the intervening variables exerted no influence on the relationship. Any discrepancy between these two variables could be attributed to the influence of the intervening variables. The approach of the present study was to analyse the independent variable and the intervening variables and draw inferences regarding the dependent variable.

The simple premise underlying the method of evaluating teacher effectiveness is as follows: if the student is intrinsically motivated, he will learn in order to satisfy his own needs and will continue learning even outside the classroom. It rests, then, with the evaluator to ascertain two things: (1) the degree of students' motivation, and (2) the source of this motivation. It is probable that there are overt manifestations of the students which would reveal de-

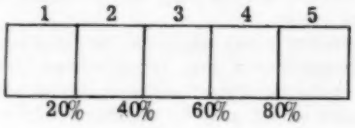
EVALUATION OF TEACHER ABILITY

I. Effect of Teacher on Class (Effect of Environment)

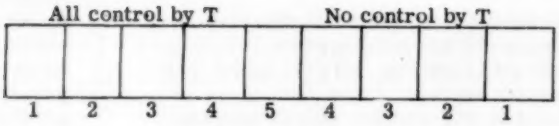
Appearance (physical and dress)
Voice
English usage
Mannerisms (absence)
General vigor and enthusiasm
Poise
Emotional stability
Competency in subject area
Physical climate



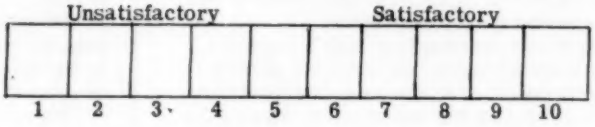
II. Reaction to environment (of class to teacher)



III. Adequacy of Student



1. Outcome of General Experience



FRAME OF REFERENCE

Scoring

I. Effect of Environment

This includes all aspects of environment set up by the teacher in the classroom.

A rating of 5 would be:

Appearance	Nothing about face or body to be a distraction from learning. Well groomed.
Voice	Well modulated, normal pitch, timbre, volume well controlled. Teacher can be heard and understood by students.

English usage	Correct grammar and correct choice of words. Pronounciation and enunciation are correct and clear.
Mannerisms	Free from distracting idiosyncracies. These would include any word or action repeatedly used.
General vigor, enthusiasm	Interest in classroom situation. Manifested by voice, posture, expression.
Poise	General demeanor in all normal situations. Confidence and ease in handling classroom situation.
Emotional stability	Manifested by teacher's reaction to emotion provoking situations. Also, emotional reaction to ordinary classroom incidents.
Competency in subject areas	Only graded when there is significant indication that the teacher is not well informed about subject.
Physical climate	Teacher controlled physical aspects of classroom conducive to learning, e.g., books, pictures, light, temperature.

Each five minutes, the classroom situation was adjudged for each category where there was evidence. If a critical incident occurred, the 5 minute period limit was disregarded.

II. Reaction to Environment (of class to teacher)

Graded according to percent of students who are apparently attentive. The students motivation may be either intrinsic or extrinsic. Here, all that is looked for is evidence of attention.

III. Adequacy of Student

Grade:

1 Teacher attempts to subjugate students to her will. She may or may not be successful.

5 The teacher leads but gives credence to student needs and interests. Highest rapport here. There is a balance of control between students and teacher.

1 Teacher allows students complete control of the situation—no leadership shown.

If there is difficulty in distinguishing between types of control, observe effect of teacher leaving classroom. If direction is maintained, the motivation is probably intrinsic.

IV. Outcome of General Experience

Add II and III.
If the student is attentive and there is a balance between student and teacher control, intrinsic motivation may be assumed. In this situation, self direction is also inferred. Individual differences would be met and there would be transfer value to other situations. The higher the sum of II and III, the greater the student's attitude for positive self direction.

gree and source of interest.

Evaluation, Evaluation Chart

After two observers had practiced observing teachers and classes during teaching and learning situations, they observed the experimental group (N = 13) six times each (each period was 20 minutes) and rated them every five minutes, (twenty-four five minute ratings on each individual). The results of the five minute ratings were added and a mean determined. The final grade was the sum of the averages of: effect of teacher on class; reaction of class to teacher; adequacy of student; and outcome of general experience.

The reliability of scoring was .963 (Pearsonian r) for a subjective device this was an extremely high reliability coefficient and may be explained by: (1) a great deal of practice before and during the evaluation of the experimental subjects; and (2) discussion between the two observers after each practice observation period to improve the frame of reference. This latter means that the later practice observations would have greater reliability than the earlier. Analysis of this factor demonstrated that such was the case to an insignificant degree (.50 level).

Though the evaluation chart was a total analysis, each of the four factors was related to the whole as shown by the results in Table I.

It is noted that, though the parts apparently share a common factor, there are also specifics.

The problem of how many times observations should be made is a significant one. One of the major criticisms which could be leveled at the proposed method of evaluating teacher effectiveness is that the observers, themselves alien to the classes, modify the situation they are observing, and that this modification prevents valid conclusions. Since there was no conceivable way of directly observing without being in the classroom it was hypothesized that becoming part of the classroom situation would eliminate grounds for this criticism. It was found that, at first, the students regard the observers with curiosity and the teachers themselves may have been influenced. Soon the teachers and students appeared to attend very little to the observers. Of course, it was impossible to measure directly the influence of the observers, but some indication was found in the observers' reports. They stated that, at first, the students showed excellent discipline which later, in some cases, became bad; first, some of the teachers used visual aids and other classic synonyms or accepted stereotypes of good teaching while later they did not; objective evidence was found in the poorer rating which all but two teachers achieved on the last three of the six evaluations. The mean of the first three observations was 16.4; the mean of the second was 13.3 (C.R. = 1.30, significant at .20 level). The relationships between the first three observations and the last three was .965 (N = 13). It was hypothesized that facility to withstand and adjust to the presence of potentially threatening stimuli in the classroom was associated with teaching effectiveness. This may be true, but the difference between the first three and the last three observations was insignificant indicating that the first three observations alone would have been adequate for judging teacher effectiveness in those teaching situations.

Problem II: The second step in the evaluation and prediction of teacher effectiveness was to utilize instruments which measure at a molar level attitudes particularly relevant to teaching effectiveness.

The instruments used in the published research reviewed included various personality tests, intelligence tests, academic records, case record data, and numerous rating scales. Many different criteria of prediction have been employed and many intercorrelations reported, but Barr states, "Where several criteria are employed to measure success in a given respect, as for example, success in service, the intercorrelations are low" (2:224). He further infers that these different criteria apparently measure different factors.

Probably the most effective method which would fulfill the criteria previously stated would be by use of projective devices. The results of previous studies utilizing projective techniques, though inconclusive, are significant to the present study. In an investigation of the value of quantitative Rorschach factors in the evaluation of teacher effectiveness, no relationship was found between pupil ratings on a standardized check list of teacher traits and personality maladjustments of selected teachers using the Monroe check list. When the Maile and Harrower-Erickson list of psychoneurotic signs was used, a .520 relationship was found between absence of neurotic signs and favorable rating (6).

A study concerning the prediction of the ways in which teachers interact with children was made which utilized eight pictures depicting situations within the experience range of teachers and related to adults' interaction with children. The study indicated the possibility of predicting teacher behavior in the classroom.

The following conclusion seems significant to the present study:

> To date, the best predictions seem to be had from combinations of so-called subjective and objective measure. While subjec-

TABLE I

INTERRELATIONSHIPS BETWEEN THE PARTS OF THE EVALUATION CHART OF TEACHER EFFECTIVENESS
(Contingency)
N = 13

	Part I	Part II	Part III
Total	.734***	.830***	.779***
Part I		.609**	.679**
Part II			.523

* Coefficients of correlation are interpreted by F test of significance. See reference 12.
** .05
***.01

tive measures are in general more unreliable than objective measures they appear to provide in some instances data on aspects on efficiency not as yet adequately covered by objective measures (1).

To test the hypothesis that teaching is a complex which must be predicted as a whole, it was concluded that the method offering greatest possibilities for investigating the qualities and quantities which constitute this whole is by use of the projective technique which makes it possible to analyse personality on a molar level. This technique assists in distinguishing between opinion and attitude and facilitates the investigation of behavior dynamics.

Problem A: The first step in the analysis of the individual in the attempt to find correlates with teacher performance was to utilize a measuring instrument which probed the potential adjustment pattern of the subject.

Degree of adjustment is dependent upon inner stability, intellectual control of emotion, and rapport with the external environment. The Rorschach is a holistic test which purportedly would reveal the potentialities of the subject to:

1. Establish rapport between himself and his internal environment, and define his level of maturity and inner stability.
2. Have intellectual control of his emotions.
3. Establish rapport between himself and his external environment.

All these factors are presumably important to the teacher—inside and outside the classroom.

To fulfill the above requirements, the group Rorschach was attempted. The method of scoring was based on seventeen signs of adjustment.

Administration

Each subject was given a booklet of ten pages. Each page had a stenciled reproduction of a Rorschach blot. At the bottom of the page was a section for recording what was seen during the test and another section to record during the inquiry.

Before actually beginning the test, the following directions were given:

> We are going to show you some ink-blots —you all know what an ink-blot is—and we would like to have you write down everything you see in the space marked 'What you saw in this blot'. People see all sorts of things in these blots and you write down everything you see. Pay no attention to the rest of the answer booklet at this time—just the lower left hand corner. Remember, there are no right or wrong answers—just write down everything the blot might be for you—everything it makes you think of.

Each blot was then projected on a screen and adequate time was allowed for all subjects to finish.

After the ten blots had been presented, the following directions were given:

> Now we are going to show you the blots again and this time we want you to do two things:
> 1. Locate where you saw each object by circling it on the stenciled blot and labeling it with the number of the object you write down. Do this for each object seen.
> 2. Under the column marked "describe what you saw. What in the blot made you think of this?" tell how you saw the object, what made you think of it—what decided you that the object looked like what you decided it did? These determiners might have been color, shading, form, texture or third dimension. If there was only one determiner, put it down, but if there were more, put the strongest or most important down first followed by the next important one, etc.

After any questions had been answered, the inquiry and the identification were completed.

The protocols were then scored according to standard directions (10).

Tabulations

The following adjustment signs were employed as the measure of potentiality for adjustment (7):

1. M ⩾ FM
2. M, 3 or more
3. ΣC > Fc + c + C'
4. F% 50 or less (not less than 20%)
5. Dd + S, 10% or less
6. P, 4 or more (not more than 30% of R)
7. R ⩾ 20
8. FC > CF
9. FC, 2 or more
10. No C
11. Percent of responses to last three cards 40%
12. FK + Fc, 2 or more
13. W:M 2:1
14. A%, 50 or less
15. No color shock
16. No shading shock
17. No refusals

The responses were scored twice and evaluated on the basis of the seventeen signs of adjustment. Since the first fifteen signs were

objectively evaluated, they alone were correlated with the criterion (observation) and the relationship was changed insignificantly. There is a possibility, then, that for the purpose of teacher prediction, fifteen signs would be adequate. For purposes of the present experiment, however, seventeen signs were used.

Problem B: The second phase of the selection of predictive tests was to devise a T. A. T. type test for analysis of attitudes particularly relevant to the problems which might be significant to teaching efficiency and to learning efficiency. The values derived from this test must necessarily be the result of potential and application, or aptitude and ability.

Subjects and Materials

The subjects were graduate and undergraduate, female and male students from a southern university.

Originally, twenty pictures were devised which depicted critical situations in areas where adjustment would be significant to the individual's teaching performance. These areas were concerned with attitudes of self adjustment and social adjustment in life situations. The twenty pictures were reduced to ten by the elimination of those which led to duplication of response, appeared trivial, were not sufficiently stimulating of projected response, were too pointed or were too emotion provoking. It is felt that the ten remaining pictures cover many of the major areas to which the individual must adjust.

The following pictures were included in the final form of the test:

	Area of Adjustment
1. Children apparently laughing in classroom situation	Social Academic
2. Girl reading while other children are playing in play-ground situation	Social (position in peer group)
3. Boy turning hose on two children, one without bathing suit	Sex Social
4. Two boys of possibly different economic classes with bat held between them	Economic Social
5. Three children around what may be a grave	Death Social
6. Two children in what might be a wrestling match	Sex Social
7. Children apparently going to church while two boys are apparently fishing	Religion Social
8. Boy playing with several girls	Sex role Social
9. Two white boys and a Negro boy playing marbles	Race Social
10. Child looking into bassinet	Family Sibling rivalry

Certain other attitudes may be expressed if the need is sufficiently strong.

The method of administering, scoring, and interpreting the T. A. T. responses is discussed.

The subjects were given the following directions:

> We are going to present some pictures to you and ask you to make up a story about each. You will react to each situation as an educator—an educational psychologist—in which you will find the most significant problem, and solve it.

The following is the frame of reference for scoring the T. A. T.:

FRAME OF REFERENCE FOR SCORING THE T. A. T.

T. A. T. is scored from two standpoints: the ability to find the problem and the ability to solve the problem once it has been found.

1. Awareness of the particular problem: There are five possible points for each picture with five being the highest score and one the lowest. The frame of reference for scoring in this category is as follows:

5 points: Awareness of the particular problem
- Picture I: Ridicule, either of another member of the class or of the teacher.
- Picture II: Security in the group
- Picture III: Awareness of the child without a suit
- Picture IV: Economic differential
- Picture V: Death
- Picture VI: Sex

Picture VII: Religion
Picture VIII: Sex role of boy
Picture IX: Race
Picture X: Sibling rivalry

4 points: Awareness of a significant problem. Modification of one of the above or another problem not in the above list but which has significance from the standpoint of the educational psychologist.
3 points: Some problem, though not a significant one is chosen.
2 points: Avoid problem. May show by answer that some problem was felt but does not indicate a real one.
1 point: Sees no problem at all.

2. Solution

5 points: Either offers an intelligent suggestion for problem solution as an educational psychologist or shows logical outcome of the problem chosen. Must get five on problem selection to get five on solution.
4 points: Same as above on a four problem.
3 points: Either logical outcome of insignificant problem or some suggestion of educational psychologist.
2 points: Story simply ends. Problem is not carried through to conclusion.
1 point: Suggests a solution to a problem which is incongruous. May have drastic outcome for very insignificant problem or may have illogical ending for significant problem.

Evaluation, T. A. T.

The only evaluation of T. A. T. was from data derived from the experimental sample (N = 13).

Reliability of scoring: When each test was scored independently by two clinicians, the reliability coefficient was found to be .9075 (N = 13).

Reliability of the test: The reliability of the test using the test-retest method with a two month interval was .911 (N = 77). This was found during the exploratory trials and indicates a high level of consistency.

Relationship to Rorschach signs: When the relationship between the T. A. T. and the seventeen Rorschach signs was determined, (N = 13), the resulting coefficient of correlation was found to be .605 (.05).

Scoring: The T. A. T. was administered to six young college professors—three members of the psychology department and three members of the education department as an index of scoring validity. This method of establishing validity of scoring may be challenged because nothing about the teaching effectiveness of psychologists or educators was known. However, specialists in the areas of psychology and education should perform highly on an instrument purportedly concerned with the measurement of understanding and application of psychological principles.

It was impossible to observe the teaching effectiveness of these six subjects, but an analysis of their test performance would indicate the validity of the foregoing hypothesis.

Internal consistency: When each picture was related to the total T. A. T. score, (N = 13), the relationships were as shown in Table II.

TABLE II

THE PARTS RELATED TO THE WHOLE
(Contingency)

Picture	r
I	.626*
II	.714**
III	.829**
IV	.723**
V	.651*
VI	.776**
VII	.562*
VIII	.914**
IX	.407
X	.899**

* .05
**.01

What the T. A. T. Tests: Being able to find the problem probably requires a certain amount of sophistication, but this ability appears more related to lack of emotional blocking. A subject, for example, might readily be able to recognize problems of a certain nature but be unable to respond to stimuli which were ultimately associated with poor adjustment areas of his own. Alertness, or sensitivity to social situations, too, is probably important here.

Solution of the problems requires training in psychology and its application. A knowledge of child and adolescent growth, development, and behavior is essential. Possibly with this knowledge certain reserves would be removed so that the individual could more readily recognize the problem in the stimulus situation and react to it in a sophisticated, mature manner.

Problem III: The problem in the third phase of the evaluation and prediction of teacher effectiveness was to ascertain the relationship between qualities possessed by teachers and teacher efficiency with the aim of predicting teacher success.

Subjects and Materials

There were thirteen volunteer subjects. They

were all teachers in the only high school open during the summer months in a southern city. There were nine men and four women teaching the following subjects: general science (2), social science (3), Latin (1), mathematics (3), English (3), and business (1).

There were thirty-two teachers who were on duty at this time. The projected experiment was explained to each prospective subject and thirteen volunteered; the others were too busy, did not want their classes disrupted by the observers, were suspicious of the tests, or were suspicious of the experiment in general. It is possible that the final thirteen were a selected sample and not representative of the entire group, but an analysis of observation and test results does not explain the basis of selection. A larger number of subjects would have yielded more valid results and the investigators made every conceivable effort to increase the sample size. The sample size of 13, however, was found in keeping with the physical facilities possessed by the present investigator.

The subjects were written letters which re-explained the purpose of the experiment and asked when an appointment for testing would be convenient. The time was arranged.

The following were the materials used:

1. The Projective Technique for the Analysis of Teacher Aptitude
2. The group Rorschach
3. Results of observations
4. Age of the subjects

Method of Procedure

Administration of the tests: At the beginning of the testing session the purpose of the experiment was again briefly explained, then the tests were administered in the following order:

1. T. A. T. It was administered first to reduce the tendency on the second test to analyse the stories as teachers, because it required the greatest intellectual alertness, it was given while the subjects were least fatigued.
2. Rorschach. Administered last so perseveration would not affect the other test.

There was a ten minute rest period between the T. A. T. and the Rorschach. It is understood that it would have been advisable for the tests to have been administered on separate days, but in the light of the situation, this was impossible.

Scoring the tests: To control the effect of contamination, each test was scored without reference to the name of the subject (each subject was given a number, varying from one set of tests to the other). No scorer of any one test was able to associate the results of that test with any other evidence (except handwriting).

The Rorschach and T. A. T. scorers were well trained for their specific tasks.

The observations: Before the actual observations of the experiment took place, two graduate students in education from a southern university were instructed in the use of the observational technique and practiced its application. After each trial, they brought their evaluations together with the project director and reviewed their results. From the first, there was little variation in their findings, and, after certain semantic problems had been solved, the experimental observations were begun.

Each subject was observed six times in twenty-minute sessions. During the twenty minutes, four five-minute observations were made. The average of the four observations was the final score for the twenty-minute period. If a critical incident occurred, the five-minute periods were not rigidly adhered to.

The school day consisted of four and one-half hours (three one and one-half hour periods), and each subject was observed longitudinally (for the week) and cross-sectionally (for the day). The method of scheduling the time of observations followed the typical latin square design.

Each subject was observed six times: three times in the early morning, and three times in the late morning and afternoon.

Results: The following data were obtained from each of the thirteen experimental subjects: age, sex, number of years' experience, subject taught, Rorschach score, and T. A. T. score.

The number of years' experience when age was held constant appeared to have little to do with teaching effectiveness (partial r = 0315). The relationship between experience and effectiveness was -. 494; age and effectiveness was -.507; age with experience was . 952.[1] Sex and subject taught too were not found to be significantly related to the criterion or the predictive measures.

Table III presents the measures of central tendency and variability of the four selected measuring devices.

It is noted from Table IV that the T. A. T. is the best individual predictor of the criterion (r^2 = .561).

The . 507 coefficient between the criterion and age was an interesting one. Slight positive relationships have been reported (12, 13), and, more recently, slight negative (14).

1. Some early studies have reported high positive relationships between "qualities of merit" and years of experience, using superintendents' and principals' ratings as criterion. See references 4, 5.

TABLE III

MEASURES OF CENTRAL TENDENCY AND VARIABILITY OF THE FOUR SELECTED MEASURING DEVICES

	M	σ_M	σ	σ_σ
Observation	15.04	1.34	4.65	.95
T.A.T. II	65.81	4.22	15.61	2.98
Rorschach	8.23	.89	3.09	.63
Age	42.04	3.11	10.76	2.20

TABLE IV

INTERRELATIONSHIPS BETWEEN THE CRITERION, THE RORSCHACH, T.A.T. II, AND AGE

	T.A.T. II	Rorschach	Age
Criterion	.749**	.608*	-.507
T.A.T. II		.643*	-.196
Rorschach			-.266

* .05
** .01

There may be a curvilinear relationship between age and teacher effectiveness. The data from the experimental sample (N = 13) did not indicate so, nor has previous research (14).

The negative relationship between age and the Rorschach adjustment indicators, though insignificant, is in the direction of previous findings (8).

Table V presents partial coefficients of correlation.

TABLE V

SIGNIFICANT PARTIAL COEFFICIENTS OF CORRELATION BETWEEN THE TESTS OF THE BATTERY

r 12.3 = .589*	r 13.4 = .560*
r 13.2 = .253	r 14.3 = .451
r 12.4 = .769**	r 13.24 = .183
r 14.2 = .555*	r 14.23 = .082
r 12.34 = .643**	

* .05
** .01

1 = Criterion
2 = T. A. T.
3 = Rorschach
4 = Age

Probably the most significant feature of the foregoing table is the relationship between teaching effectiveness and age. Evidently the age concommitants are the important thing and, if the teacher can retain youthful characteristics in matters of alertness and mental health, age, per se, will have little bearing on her effectiveness as a teacher. It is only insofar as these concommitants appear a constant, linearly related with age, that age is useful in predicting teacher effectiveness.

The multiple coefficient of correlation was determined to find the optimum weighting for combining the variables to predict the criterion. Table VI presents the significant conclusions derived in ascertaining the multiple correlation corrected for shrinkage.

It may be observed from the foregoing table that the inclusion of age increases $\overline{R}$; but the inclusion of the Rorschach, though increasing $\overline{R}$ slightly[2], reduces its significance. The coefficient of forecasting efficiency (E) demonstrates that the multiple coefficient is 42-43% efficient in predicting criterion scores.

The $\overline{R}^2$ indicates that 67% of the variance in the dependent variable may be attributed to variation in the three independent variables. Specific analysis using beta weights reveals that the contribution of the T. A. T. to the criterion variation is 44%; the contribution of age is 15%; and the contribution of the Rorschach is 8%. The remaining 33% of the variance of the criterion must be attributed to factors not considered in this research.

The regression equation using beta weights was:
C = T. A. T. (.598) - Age (.298) + Rorschach (.144)
The regression equation in score form was:
C = T. A. T. (.190) - Age (.129) + Rorschach (.216) + 6.16
(The 6.16 is a constant which, if the Rorschach were not included, would be 7.94)

To demonstrate the efficiency with which the regression equation predicts the criterion scores, the predictions are presented at the top of the next page.

It is seen from these predictions the degree to which the instrument is accurate and that there is little improvement in prediction when the Rorschach is included in the test battery.

Summary

Problem: The purpose of this experiment was to develop an instrument to predict teacher success.

Subjects and Materials:

Subjects:

The subjects were composed of thirteen volunteer teachers from the only high school open during the summer months in a southern city.

Materials:

1. The Projective Technique for the Analysis of Teacher Aptitude
2. The group Rorschach
3. Observational results
4. Age of the subjects

Method of Procedure:

1. Each teacher was observed six times in twenty minute sessions over a seven week period. The observations were so scheduled that a cross-sectional, longitudinal analysis could be made.
2. The following tests were administered:
   a. The Projective Technique for the Analysis of Teacher Aptitude.

2. Insignificant increase by F test.

Subject	Actual Score	Predicted Score	
		With Rorschach	Without Rorschach
1	10.3	8.7	8.6
2	21.6	17.7	17.5
3	11.1	12.3	12.4
4	13.3	15.4	15.1
5	20.0	24.5	23.2
6	13.9	12.8	13.5
7	15.0	15.4	14.5
8	17.0	14.4	14.5
9	24.5	20.5	19.9
10	11.6	14.8	14.6
11	8.6	9.1	10.3
12	10.4	11.9	13.0
13	14.7	17.6	16.2

TABLE VI

THE EFFECT ON $\overline{R}$ OF ADDED VARIABLES

	$\overline{R}^2$	$\overline{R}$	F/.01	Significance	E	$\sigma_{est.}$
T. A. T.	.561	.749				
Age	.659	.812	9.66/7.56	.01	42%	2.71
Rorschach	.671	.819	6.11/6.99	.02	43%	2.67

b. The group Rorschach.
3. The age of each subject was obtained.
4. The test and observational results were scored and evaluated.

Results:

Though T. A. T. II was the best individual predictor of the criterion, its inclusion with age and Rorschach yielded a multiple $\overline{R}$ of .819 (σ_{est} = 2.676).

The regression equation using beta weights was:
C = .598 (T. A. T.) - .298 (Age + .144—Rorschach)
The regression equation in score form was:
C = .190 (T. A. T.) - .129 (Age) + .216 (Rorschach) + 6.16.
(6.16 becomes 7.94 if the Rorschach is not included.)

Conclusions

With reference to the foregoing data, the following conclusions may be made:

1. The method of observation appears to be a valuable tool for the estimation of teaching effectiveness. The first part of the observation chart was devoted to those aspects of the teacher which would be significant in her making an adjustment to extra-classroom society; the rest of the chart was of particular importance to the analysis of classroom effectiveness.
2. The T. A. T. appears to be a valuable tool for the prediction of significant aspects of teaching effectiveness.
3. Age is not significant in itself but certain of its residual features are. These appear to be so associated with age that age becomes a significant element in the prediction study.
4. The Rorschach appears not to add significantly to the accuracy of prediction of teaching success. It is an important measure, however, because it gives the adjustment potential inside and outside the classroom; and since teacher success is considerably more than teaching success, this potential must be considered. It is suggested, then, that the Rorschach need not be included in the regression coefficient but that a minimum of nine be obtained from the seventeen adjustment signs for teachers and for prospective teachers.
5. The value of this prediction device may be found in the promotion and selection of teachers in the field. Further study is being made to determine its value in the selection of teacher candidates.
6. The conclusions presented here appear provocative and warrant the repetition of the experiment with a larger experimental population in a situation possessing greater facilities (money, trained assistantships, and time).

BIBLIOGRAPHY

1. Alexander, Theron. "The Prediction of Teacher-Pupil Interaction with a Projective Test," Journal of Clinical Psychology, VI (July 1950), pp. 273-276.
2. Barr, A. S. "Teaching Competencies," in Encyclopaedia of Educational Research (New York: Macmillan Co., 1950).
3. ———. "The Measurement and Prediction of Teaching Efficiency: A Summary of Investigations," Journal of Experimental Education, XVI (June 1948).
4. Boyce, A. C. "Methods of Measuring Teachers' Efficiency," in Fourteenth Yearbook of National Society for the Study of Education, Part I (Bloomington, Ill.: Public School Publishing Co., 1915).
5. ———. "Qualities of Merit in Secondary School Teachers," Journal of Educational Psychology, III (1912).
6. Cooper, J. G. and Lewis, Roland B. "Quantitative Rorschach Factors and the Evaluation of Teacher Effectiveness," Journal of Educational Research, XLIV (May 1951), pp. 703-707.
7. Davidson, Helen H. "A Measure of Adjustment Obtained from the Rorschach Protocol," Journal of Projective Techniques, XIV (March-December 1950), pp. 31-38.
8. ——— and Kruglov, Lorraine. "Personality Characteristics of the Institutionalized Aged," Journal of Consulting Psychology, XVI (1952).
9. Johnson, G. B. "The Origin and Development of the Spanish Attitude Toward the Anglo and the Anglo Attitude Toward the Spanish," Journal of Educational Psychology (November 1950).
10. Klopfer, Bruno, and Kelly, Douglas. The Rorschach Technique: A Manual for a Projective Method of Personality Diagnosis (New York: World Book Co., 1942).
11. Knight, F. B. Qualities Related to Success in Teaching, Contributions to Education, No. 120 (New York: Teachers College, Columbia University, 1922).
12. McNemar, Quinn. Psychological Statistics (New York: John Wiley and Sons, 1949).
13. Odenweller, A. L. Predicting the Quality of Teaching, Contributions to Education, (New York: Teachers College, Columbia University, 1936).
14. Reed, Harold J. "An Investigation of the Relationship Between Teaching Effectiveness and the Teacher's Attitude of Accept-

ance," Journal of Experimental Education, XXI (1953).

15. Rotter, Julian R. "Thematic Apperception," Journal of Personality, XV (1946).

ATTITUDES AND PROFESSIONAL RELATIONSHIPS OF TEACHERS: A STUDY OF TEACHER MORALE*

E. C. HUNTER
Tulane University
New Orleans, Louisiana

Introduction

INDUSTRY, business and education are increasingly recognizing the significance of human factors. Many studies indicate that social factors are the most important determinants of productivity and success in human enterprises. The major task of any organization appears to be the creation and continuance of a favorable social and emotional climate that will capitalize on the potentialities of workers and provide the basic satisfactions that people want. In situations where workers obtain these satisfactions, attitudes of interest, cheerfulness, initiative, devotion, and cooperation are evidenced. Increase in production and growth in unity, strength and effectiveness usually accompany such attitudes. The mental and emotional reactions to the job imply high morale by which we mean "the capacity of a group of people to pull together persistently and consistently in pursuit of a common purpose" (7) or, "the quality of giving fully of one's best efforts to carry out a purpose" (1) or, "the capacity and courage to carry on a task with determination, loyalty, cooperation, and a sense of personal satisfaction and well-being" (6). On the other hand, low morale is indicated by loafing, bickering, absenteeism, dissatisfaction, lack of confidence, high turn-over, inefficiency, and low production. It is essential that management and workers agree on and pursue common goals, but perhaps the most important element in morale is what the workers feel and believe, not what actual conditions are.

Morale can be sampled and its quality determined by studying the attitudes and reactions of workers to their jobs, the satisfactions, dissatisfactions, and feelings being experienced. Attitude surveys to discover how employees react to a variety of practices and policies have been used extensively in business and industry. In a few school systems, surveys have been conducted in which opinions of teachers were sought on matters relating to their jobs. Keeping alert to the state of mind of teachers with respect to changes in school practices would seem to be necessary for maintaining a good working atmosphere and morale. Periodic attitude surveys stimulate morale in that teachers are led to believe that somebody is interested enough to try to find out how they feel. In some school systems, annual surveys are being conducted to identify teacher attitudes toward important factors affecting morale, the results being used to identify and improve conditions adverse to morale.

Selected Related Studies

Major factors identified by Chase (3) as being important for job satisfaction in teaching include freedom of the teacher to plan his own work, salary, quality of professional leadership and supervision, opportunity for teachers to participate in educational planning and policy making, and adequacy of physical facilities. After a survey of a city school system, reported by Campbell (2), the survey staff concluded that low morale was handicapping the educational program and suggested basic points of view about which there was much confusion, followed by ways of resolving differences. Some of the confusions and dilemmas pertained to directives vs. group decisions, academic standards vs. pupil growth, purposeful learning vs. prescribed tasks, supervision vs. special teachers, staff decisions vs. anarchy, and course tinkering vs. basic curriculum study. Schultz (13), studying job satisfactions of recent University of Illinois graduates who are in teaching positions, found administrative practices and procedures, salaries, working conditions, and supplies and equipment to be closely related to morale. The superintendent was the focal point of teachers' satisfaction or disillusionment in their work. Least satisfied teachers were married men. Increased family and financial responsibility resulted in greater teacher dissatisfaction.

Hedlund and Brown (5) reported a study in which questionnaires were mailed to the home addresses of some 3000 teachers in 117 schools selected to be representative of New York State, exclusive of New York City. Questionnaires were answered anonymously and returned directly to the Division of Research, State Education Department. Findings reported were based on 1067 usuable returns. Of 75 living and working conditions, 14 were found to be critical in causing teacher turnover, an indication of low morale. These 14 conditions pertained to insuf-

* This study was made possible by a grant from the Tulane University Research Council.

ficient salary; teaching conditions such as large classes, great variations of pupils, tiring nature of teaching, and insufficient relief from pupil contact during the day; community conditions including inadequate transportation for teachers without cars, and inadequate community support for schools; administration conditions including uncertainty as to job, inadequate opportunity for professional advancement, and unsatisfactory support in handling disciplinary situations. The study further revealed that 24 to 59 percent of teachers returning questionnaires thought these conditions existed and nine to 19 percent expressed the desire to leave their jobs because of these conditions. It would appear that these results furnish lines of attack for school administrators who are concerned with improving teacher morale and teaching service.

Shilland (14) distributed a checklist of 30 items to 429 teachers in one county. On the basis of 216 returns the 10 items identified as most important to morale and arranged in order are:

1. Doing work for which one is prepared and interested
2. Adequacy of equipment and supplies
3. Consideration and courtesy by superiors
4. Physical working conditions
5. Job security
6. Administrative cooperation and assistance
7. Friendly attitude of fellow teachers
8. Fair compensation
9. Development of personality in association with and inspiring young people
10. Pupil attitudes of respect toward teaching

In lighter vein Connor (4) offered the following as satisfactions that keep a teacher on the job: a good night's sleep and waking up without an alarm clock, time for breakfast and for planning the day, good transportation, a bulletin board with only important notices on it condensed into a few words, approval from others, being recognized and evaluated by the administration, knowing children individually and seeing them grow and develop, single salary schedule, and knowing that one has accomplished something that can be built upon tomorrow.

Leipold and Yarbrough (8) reported that 1600 school people thought the most important factors affecting morale in order of relative importance are: administrative support to teachers in discipline problems, deep-seated belief in and personal enjoyment of teaching, just and adequate salary plan, proper student courtesy and respect for teachers, a good retirement and pension plan, professional attitude shown by all in handling teacher grievances, adequate sick and emergency leave policies, personal interest and confidence in ability and integrity of staff shown by administrators, cooperative spirit among teachers, security through sound tenure, and constructive and democratic supervision procedures.

According to March (11), all groups, the superintendent and board of education, the principal and his assistants, the instructional staff, the student body, and the parents, were held to be responsible for good school morale. Misunderstandings between groups were thought to be the chief destroyer of morale. To prevent this, it was pointed out, everyone must be kept as fully informed as possible as to purposes, procedures, policies, and plans of the school program. Talking about the principal and criticizing each other were given as besetting sins in teacher morale. The author agreed that anything to discourage destructive criticism would improve morale.

Many studies have shown the advantages of the group process on morale and efficiency. Typical is that by Lewin, Lippett and White (9) in which it was demonstrated that in a democratic group where policies are a matter of group discussion and decision, solidarity and identification with group goals and group products are much greater than in an autocratic group. Lippett (10) showed the effectiveness of group decisions on the rise of productivity of workers when the work group met with a skilled discussion leader to discuss work problems and to set up a group production goal. It was further found that work dropped less in efficiency, recovered faster, and showed progress to a higher level than before as a result of democratic participation in deciding on changes in job methods.

A conclusion by Wiles (15) from experiments in interviewing workers in the well-known Western Electric studies of employee motivation, reported by Roethlisberger and Dickinson (12), adequately expresses the importance now being attached to feeling elements and the group process.

> "The behavior of workers cannot be understood apart from their feelings or sentiments. Sentiments are not easily distinguished and are therefore difficult to recognize and study. Manifestations cannot be understood by themselves but only in terms of the total life situation of the person. The meaning that any individual assigns to a particular change depends on his social conditioning or the sentiments he brings to the work situation because of previous group associations, and the kind of human satisfaction he is deriving from his social participation with other workers and supervisors in his immediate group."

Purpose and Procedure

The purpose of this study is to summarize an attitudes survey of public school teachers in a large southern city. Attitudes of 968 teachers toward conditions and factors known to affect morale will be treated. Favorable and unfavorable reactions of teachers toward school and living-working conditions were investigated in 1950 and in 1953. Efforts were made to determine how teachers felt about school practices and social and professional relationships. An important phase of the study pertains to changes in attitudes over a three-year period. At the outset it was hoped that conditions found to be causing low teacher morale might be more clearly identified and appropriate action taken to remove such conditions so that a climate conducive to better teaching and learning might be provided.

The survey instrument comprised 43 "Yes" and "No" statements, drawn up in accordance with recognized technique, and designed specifically to ascertain the state of teacher attitudes and morale in New Orleans. The items were carefully selected and constructed to cover the important factors affecting morale as determined by the research in the field.

In May, 1950, the questionnaire was sent to all classroom teachers in the public schools of the city, a total of 1837 at that time. The questionnaires together with self-addressed business-reply envelopes were mailed to the home addresses of teachers who were asked to fill out anymously and check under the "Yes" and "No" columns. In June follow-up postals were sent to the teachers, urging those who had not done so to answer and return the questionnaires. By the end of July, a total of 465 completed forms had been sent back, a 25 percent return.

In May, 1953, the questionnaire was revised slightly and again mailed to all public school teachers in New Orleans, a total of 2082. After a postal reminder in June, a total of 503 completed forms were received, a return this time of 24 percent. In addition to the survey items, both questionnaires called for age, sex, race, marital status, grade or subject taught, years of teaching experience, and educational level attained. Although the total percentage return was not large, it was ascertained that returns on the basis of race, sex, education, level of teaching, and teaching experience were representative in both years.

Results

Table I shows percentages of the teacher samples for 1950 and 1953 responding to the questionnaire in the affirmative, negative and uncertain manner. The Table reveals that 75 percent or more of the teacher samples indicated favorable attitudes and high morale on fifteen items. Three-fourths of the 1953 teacher sample indicated a belief that the public schools had improved during the preceeding three-year period. More than 95 percent of the 1950 teacher sample expressed faith in the contribution of public education and indicated their belief in teaching as a profession. Approximately 90 percent of both teacher samples indicated that their teaching competence was increasing. More than 90 percent of the first sample did not find teaching monotonous and dreary. From 82 to 85 percent of both samples did not feel that they were unduly restricted in their personal and social life as a result of being teachers. These results are somewhat at variance with those reported by Hedlund and Brown (5) for New York State. More than 50 percent of the New York teachers found teaching tiring to the extent that social life after school was affected. From 80 to 83 percent of the New Orleans teachers believed their pupils showed them proper consideration, courtesy and respect. 90 percent indicated a friendly and cooperative attitude with fellow-teachers. More than 80 percent of both samples felt that they enjoyed the confidence, respect and moral support of parents and citizens. Seventy-five percent of the 1953 sample approved the junior high organization which had been in operation one year at that time.

On seventeen items from 50 to 75 percent of the teacher samples indicated favorable attitudes. These items dealt with job security, maintaining high professional standards, participation in social life and in professional organizations, improvement in teaching skill, participation in planning, attending study groups, meeting pupil needs, encouraging promising young people to enter teaching, desirable working conditions, the retirement system, cooperation with principal, recognition by parents freedom to speak out on issues, and administrative and supervisory assistance. Hedlund and Brown (5) found that 33 percent of New York State teachers felt that working conditions did not encourage the teacher to improve upon the quality of teaching. This figure was slightly more than 50 percent for the New Orleans teachers.

On the other hand, Table I reveals that 50 percent or more of New Orleans teachers sampled indicated unfavorable attitudes on fourteen items. Seventy-seven percent of the first sample did not feel that they were rewarded in any way for exceptional services to the school or community. Approximately 60 percent felt they had an excessive teaching load. This is to be compared with 34 percent in the New York study (5). More than 50 percent indicated that their

TABLE I

PERCENTAGES OF 1950 AND 1953 TEACHER SAMPLES OF PUBLIC SCHOOL TEACHERS IN NEW ORLEANS RESPONDING TO QUESTIONNAIRE ITEMS IN EACH OF THREE WAYS: "YES", "NO", AND "?" ("?" DENOTES DOUBTFUL, UNCERTAIN, OR OMITTED)

	Questionnaire Items	1950 (N=465)			1953 (N=503)		
		Yes	No	?	Yes	No	?
*1.	Do you think the public schools in this city have improved during the past three years?				74	23	3
**1.	Do you have faith in the contribution of public education for the advancement of democratic society	96	3	1			
*2.	Do you believe the public schools are meeting adequately the needs of most of the pupils?				49	48	3
**2.	Is teaching for you a profession and not a make-shift job leading to other occupations or to marriage?	96	4	0			
3.	In your opinion, are high professional standards in education being upheld in this city and state?	44	52	4	65	31	4
4.	Do you encourage promising young people to enter the teaching profession?	66	32	2	69	28	3
5.	Do you feel you are doing the work for which you are best prepared?	93	6	1	89	10	1
6.	Do you believe that your competence as a teacher is increasing?	89	8	3	91	7	2
7.	Do you think that teaching imposes severe strain on your physical and mental endurance?	49	48	3	48	49	3
**7a.	Do you find teaching monotonous and dreary?	6	93	1			
8.	Have you been requested to teach subjects or grades in which you are not prepared or interested?	40	59	1	26	74	0
9.	Do you have adequate training for teaching mentally backward and mentally gifted children?	48	48	4	34	62	4
*10.	Do you believe that the primary aim of effective discipline is to teach obedience and conformity and to secure good order?				48	48	4
11.	Are you unduly restricted in your personal and social life as a result of being a teacher?	17	82	1	14	85	1
12.	Are your general working conditions such as to encourage you to give your best services to classroom teaching?	45	54	1	50	49	1
13.	Do you have an adequate amount and proper kinds of equipment and supplies for good teaching?	39	60	1	41	58	1
14.	Do you feel you have an excessive teaching load at present?	39	60	1	41	59	0
15.	Is your salary a fair compensation for your services?	37	62	1	43	54	3
16.	Are you able to maintain the standard of living which good teaching requires with reasonable freedom from economic worry?	42	55	3	43	50	7
17.	Is the retirement system (City or State: Underscore which) satisfactory to you?	58	34	8	63	25	12
18.	Do you think you have adequate protection for yourself and dependents in case of sickness or accident?	42	56	2	48	45	7
19.	Is it necessary for you to accept outside work during the school year or during vacation periods to provide yourself an adequate income?	47	49	4	47	45	8
20.	Do your pupils show you proper consideration, courtesy and respect?	83	15	2	80	13	7
21.	Is there a friendly and cooperative attitude between you and your fellow-teachers?	90	9	1	91	4	5
22.	Do you enjoy the confidence, respect and moral support of parents and citizens in your school community?	83	10	7	84	8	8
23.	Are you recognized and rewarded for exceptional services to the school and community?	18	77	5	42	45	13

TABLE I (Continued)

	Questionnaire Items	1950 (N=465) Yes	No	?	1953 (N=503) Yes	No	?
24.	Do you often seek counsel and advice from your principal?	53	45	2	59	34	7
25.	Do you participate responsibly in the planning of the program and activities for your school?	67	31	2	72	21	7
26.	Is the quality of supervision in your school constructive and conducive to improvement in teaching?	51	46	3	57	36	7
27.	Is your work properly evaluated by the administrative and supervisory personnel?	39	54	7	46	42	12
28.	Are you given proper recognition for your work by parents?	59	33	8	67	22	11
29.	Are you encouraged to grow in proficiency and increase your skill as a teacher?	66	31	3	68	25	7
30.	Does the system provide satisfactorily for advancement and promotion on merit?	17	76	7	25	61	14
31.	Does the security of your job depend on your remaining silent concerning important issues with which the administration is faced at times?	31	63	6	24	63	13
32.	Does the administration and supervisory staff assist you and cooperate with you in every way possible in the performance of your duties as a classroom teacher?	41	54	5	50	39	11
33.	Are you free to pursue any course of action which you believe to be beneficial to you and your pupils?	48	48	4	50	46	4
34.	Do you participate actively in the social affairs of your school community?	70	27	3	73	25	2
35.	Are you an active participant in at least one professional organization?	79	20	1	79	20	1
36.	Are you encouraged and aided by the school system to attend workshops and study groups for the purpose of studying professional problems?	43	54	3	87	12	1
37.	Do you visit in the homes of most of your pupils?	13	86	1	17	81	2
38.	Are you encouraged to make constructive suggestions relative to school policies and practices?	38	59	3	55	42	3
39.	Do you feel there is wholesome cooperation between the public schools on the one hand and the non-public schools on the other?	41	54	5	55	38	7
40.	As a teacher, do you derive inspiration and encouragement in the personal and professional relationships in this school system?	41	54	5	53	41	6
*41.	Do you approve the junior high school organizational plan as recently instituted in this city?				75	20	5
*42.	Do you think the high schools should provide more definite job training for many students?				86	11	3
*43.	Do you believe that teachers should have opportunities to qualify for salary increments through exceptional service to pupils, school, and community?				61	32	7

* Not included in 1950 questionnaire.
** Not included in 1953 questionnaire.

work was not properly evaluated by administrative and supervisory personnel. From 61 to 76 percent did not believe the system provided satisfactorily for advancement and promotion on merit. In the New York (5) survey, 37 percent of the 1067 teachers sampled felt that the opportunities for advancement were inadequate. In the present study, from 81 to 86 percent indicated that they did not visit in the homes of most of their pupils. In 1950, more than 50 percent did not feel there was wholesome cooperation between the public and the non-public schools. More than 50 percent of the 1950 sample indicated they did not derive inspiration and encouragement from working in the school system. From 54 to 62 percent did not agree that their salary was a fair compensation for their services. In the New York study (5), 51 percent believed that salary was insufficient. Sixty-two percent of the second New Orleans sample did not feel that they had adequate training for dealing with atypical pupils. Approximately 60 percent in both years did not feel they had adequate equipment and supplies for good teaching.

Table I further shows differences in percentages of teachers indicating favorable attitudes and good teacher morale for the 1950 and 1953 teacher samples on 35 items. The 1953 sample showed gains in percentages on 29 items over the 1950 sample. The gains were more than 10 percent on eight items. On six items percentage losses were sustained indicating that a smaller proportion of teachers showed favorable attitudes on these items in 1953 than in 1950. However, on only one of the six items, number nine, was the loss as much as 10 percent. On two items no percentage changes were noted for 1950 and 1953. The remaining nine items were not included in both questionnaires and thus could not be studied for changes over the three-year period.

Twenty-one percent more of the 1953 than the 1950 teacher sample expressed belief that high professional standards in education were being upheld. This difference is significant beyond the one percent level of confidence. Forty percent of the 1950 sample and 26 percent of the 1953 sample indicated that they had been requested to teach subjects or grades in which they were neither prepared nor interested. This 14 percent decrease is a significant difference at the two percent level and reflects a definite improvement in personnel policy and teacher morale over the three-year period. On item 23, "Are you recognized and rewarded for exceptional services to the school and community?", eighteen and 42 percent of the 1950 and 1953 samples respectively, responded in the affirmative. This percentage difference of 24 points is significant beyond the one percent level. (Part of this gain may be due to the fact that the item in the 1950 questionnaire did not contain the word "recognized".) On item 36, "Are you encouraged and aided by the school system to attend workshops and study groups for the purpose of studying professional problems?", forty-three and 87 percent of the 1950 and 1953 samples respectively, checked the favorable "Yes". This difference is significant beyond the one percent level. (Results of this item for the two questionnaires are not directly comparable since the statement in 1950 referred only to workshops and study groups "outside the city".)

In 1950 only 38 percent of the sample indicated that they had been encouraged to make constructive suggestions relative to school policies and practices. In 1953, this percentage had increased to 55, a gain of 17 points which is significant at the two percent level. Fourteen percent more of the 1953 than the 1950 sample indicated they felt a wholesome cooperation between the public and non-public schools. This difference is significant at the five percent level of confidence. From 41 to 53 percent of the 1950 and 1953 samples indicated that they had derived inspiration from their personal and professional relationship with the school system. This gain of 12 percent is appreciable but is a significant difference not beyond the ten percent level of confidence.

On item 9, forty-eight percent of the 1950 sample and 34 percent of the 1953 sample indicated they had adequate training for handling exceptional children, the backward and gifted. This difference is significant at the five percent level in favor of the 1950 teachers. It is probable that the 1953 teachers had become better informed about wide ability ranges and consequently were better able to appreciate their limitations in dealing with variants than the 1950 teachers.

Summary

1. Relatively high percentages of teachers in this study, 75 percent or more, indicated favorable attitudes toward items dealing with personal adjustment to teaching, faith in the contribution of education, increasing teacher competence, pupil-teacher-parent relationships, the junior high school organization, and more adequate job training for high schools.
2. Better than average percentages of the teachers, 50 to 75 percent showed favorable attitudes toward questions dealing with job security, retirement, educational standards, meeting pupils needs, recruiting promising young people, participation in social life, working conditions, administrative and supervisory assistance and cooperation, freedom to teach and to speak out on issues, participation in planning, and cooperation between public and non-public

schools.

3. Less than half the teacher sample, 13 to 50 percent, showed favorable attitudes toward items dealing with teacher competence in handling atypical pupils, wholesome school discipline, adequacy of supplies and equipment, teaching load, financial security, recognition and reward for exceptional services, having work properly evaluated, advancement and promotion on merit, and visitation in the homes of the pupils. Unfavorable attitudes and low morale on these items were indicated by 50 to 86 percent of the teachers in this study.
4. Marked and significant increases in percentages of the 1953 teacher sample over the 1950 sample, indicating improvement in attitudes and morale over the three-year period, were made on questions dealing with maintenance of high professional standards, assignment of teachers, recognition and reward for exceptional services, being aided and encouraged to attend workshops, being encouraged to make constructive suggestions relative to school policy and practice, cooperation between public and non-public schools, and deriving inspiration and encouragement from working in the system.

Implications

The evidence from this study clearly suggests that definite progress has been made in the New Orleans Public School System in developing constructive and cooperative relations between classroom teachers and administrative officials. On 29 out of 37 items a larger proportion of the 1953 than the 1950 teacher sample indicated favorable attitudes. Although the morale trend is upward, results on practically all items leave room for more improvement. Considerably larger proportions of teachers than is true at present, apparently, should:

1. Encourage able young people to enter the teaching profession
2. Be helped and encouraged to acquire specific training for dealing more effectively with atypical pupils
3. Feel that they are working under favorable conditions
4. Be supplied with adequate materials
5. Feel that they do not have excessive teaching loads
6. Be able to maintain a decent standard of living with reasonable freedom from worry
7. Be recognized and rewarded for exceptional services
8. Work cooperatively and constructively with administrative officials
9. Be helped to develop a better understanding of discipline
10. Feel that they are becoming better teachers
11. Have favorable attitudes toward advancement on merit
12. Participate more actively in the life of the community
13. Visit in the homes of more pupils
14. Harbor favorable attitudes toward cooperation between public and non-public schools
15. Experience inspiration from working in the system

Teacher morale is influenced by many factors and conditions. Teachers, the superintendent and board of education, the administrative and supervisory staffs, and the community should share in the responsibility for keeping morale high. Teachers can accomplish much for themselves on their own initiative through personal and professional growth and continued educational and cultural advancement. The community and school administration can do much by way of increasing support of schools, providing better working conditions for teachers, and developing a high level of cooperative leadership throughout the school system.

REFERENCES

1. American Association of School Administrators, Morale For A Free World, Twenty-second Yearbook, (Washington D. C.: American Association of School Administrators, 1944) pp. 24-25.
2. Campbell, R. F., "Treating Schizophrenia to Improve Staff Moral" The School Executive, LXXIII (1954), 52-55.
3. Chase, F. S., "Factors for Satisfaction in Teaching", Phi Delta Kappan, XXXIII (1951), 127.
4. Connor, I., "Satisfactions That Keep a Teacher on the Job", Childhood Education, XXVIII (May 1952), 396-398.
5. Hedlund, P. A. and Brown, F. S., "Conditions that Lower Teacher Morale", Nation's Schools, XLVIII, No. 3, (September 1951), 40-42.
6. Katz, B. and Lehner, G. F. J., Mental Hygiene in Modern Living (New York: The Ronald Press Co., 1953), p. 434.
7. Leighton, A., "Applied Science of Human Relations", Personnel Administration, (July 1947), 4-6.
8. Leipold, L. E. and Yarbrough, J. W., "What 1600 School People Think About Teacher Morale", American School Board Journal, CXVIV (December 1949), 29-30.
9. Lewin, K., Lippit, R., and White, R. K., "Patterns of Aggressive Behavior in Ex-

perimentally Created 'Social Climates'," Journal of Social Psychology, X (May 1939), 271-299.

10. Lippitt, R. "A Program of Experimentation of Group Functioning and Productivity," Current Trends in Social Psychology, (Pittsburgh: University of Pittsburgh Press, 1948), pp. 18, 20, 21.
11. March, L. S. "Who is Responsible for Good School Morale," American School Board Journal, CXXVI (February 1953), 27-29.
12. Roethlesberger, F. J., and Dickinson, W. J., Management and the Worker, (Cambridge: Harvard University Press, 1939).
13. Schultz, R. E., "Keeping Up Teacher Morale", Nation's Schools, L, No. 4, (October 1952), 53-56.
14. Shilland, P. D., "Teacher Morale Survey," Educational Forum, XIII, (May 1949), 479-286.
15. Wiles, K., Supervision for Better Schools, (New York: Prentice-Hall, Inc., 1950), 308.

THE INFLUENCE OF A CHILD PURPOSE UPON THE PERSEVERANCE OF YOUNG CHILDREN

LOUIS ADA WILSON
Wisconsin State College
Superior, Wisconsin

Introduction

"BY PERSEVERANCE the snail reaches the ark."

"Even the woodpecker owes his success to the fact that he uses his head and keeps pecking away until he finishes the job he started."

Perseverance is a desirable characteristic and the classroom must offer frequent opportunities and provide incentives for children to develop this characteristic, to make it a habit.

During seasonal festivities children pursue activities with unusual perseverance, making gifts, planning programs, or going out for "tricks or treats". However, even when play is only moving supplies from one place to another, unloading, reloading, then returning burden to the original location, they work steadily, persevere. "Time flies" for children whenever their academic work is to result in a play, a radio skit, entertaining parents, making money for a prized objective, because a purpose pervades these situations.

All this we know from casual observation. Controlled research must support the observation and provide data on the subject.

Review of Literature

Research shows that complexity of play material, intelligence, and interest, function to sustain children's attention (Schacter 9). When children grow older the span of attention increases, but the age of the child will not determine sustained attention on an arbitrarily selected toy or activity (Herring and Koch 6). Variation of time spent with the same toy on different occasions occurs, as does decreased interest on successive opportunities with the toy (Wolfe 11). The time spent with a toy varies within each age level. (Bott 2, Van Alstyne 10, and Getteridge 5).

Could perseverance be an innate quality, just as some psychologists believe from their studies that resistance is an innate quality? (Goodenough 4 and Levy 7).

Research seems to indicate that children react to objective playthings in varied lengths of time within the same age level; that sustained attention increases with age; that age will not sustain interest in an arbitrarily selected situation; that resistance may be an innate quality; and that often students do not carry through with plans of future interest to them. A recent study shows that of 166 twelfth-grade boys in a single high school only 105 completed with help their project undertaken in the interest of occupations (Recktenwald 8).

The Problem

The investigation herewith reported resulted from an exploratory study designed to discover the effect of a "child purpose" upon perserverance of nursery school children.

The means for ascertaining the degree of perserverance were:

I. To study problem solving of the child by using objective materials. The tries were determined by counting the number of different manipulations made until the correct manipulation was used to operate the toy.

II. To study problem solving of the abstract, that which is comparable to academic or school study. Here the "tries" were determined by counting the number of answers given to each story problem presented.

III. In the practical, objective problems, the "child purpose" was an opportunity to play with tempting, intriguing, unfamiliar toys. In the abstract problems the "child purpose" was an opportunity to play with toys previously enjoyed after solving an abstract problem presented in simple story form.

Organization of the Tests

The tests were organized as follows: The Play-A-Game test series consisted of fifteen manipulatory toys. The toys required such skills as dropping objects through various sized holes; assembling objects to form a unit of play; pushing and pulling buttons to make the toys operate; pumping and tapping to operate the toy; tying and untying in order to get the toys apart and together; and pounding objects one after another into cylinder-like holes for the toy to function.

The toys were presented to the child in the order of their difficulty. Ten tests were planned, but only seven were used since all ten were

too time consuming. Each test had a score of six, except for test II of the Play-A-Game test series, which had a score of four. Perfect score for all seven tests was 40. The scores represented the number of operations needed to make the toy function. For example, in one of these tests a Pump-A-Ball toy was placed in front of the child. He was told: "You'll like this; the balls and the box go together; you can play with them; put them together." The child was given a tally for each try made in attempting to find a solution to the operation of the toy. When the test was over and scores and tallies complete, the child played with the toy.

Each of the Tell-Me test series was given a score of six. The number was arbitrarily chosen as the possible number of answers a child might make before he would become tired. The scores were added to the tallies for tries; a right answer was also a try. However, if a child persisted, his tries were far beyond the count of six A perfect total score was 36.

Terms used in the Tell-Me test series are defined: Abstraction, a word considered apart from any application to a particular object, such as, "hello"; percept, a meaningful impression of an object, dish; concept, a thought, an opinion, Dorothy is industrious or Elizabeth is cordial.

Each of the first five tests was made to present a concept which was developed from percepts presented in two sentences. Then the word, which represented the concept formed by use of the percepts, was presented in a statement. Last the child was asked a question containing the word that represented the concept which had been developed from the two percepts. The following is an example of one of this group of tests:

Dorothy works to wash the dishes.
Dorothy works to clean her shoes.
Dorothy is an industrious girl.
What do you do to be an industrious _____?

The word boy or girl was used when asking the question, as needed in the situation. The word "can" was substituted when a child said, "I don't do that." The substitution was effective in most instances.

Tests VI and VII of the Tell-Me test series were made to present a concept developed from abstract words. Then the word in each instance, which attested the meaning of the concept, was used in a statement. Lastly the child was asked a question which contained a word denoting the concept derived from the use of two abstract words. The following is an example of one of the tests:

Elizabeth says, "Hello to Mary."
Elizabeth says, "Good-bye to Mary."
Elizabeth is a cordial girl.
What do you do to be a cordial_____?

The two beginning statements of each of these tests contained words which a pre-school child could understand. The word "cordial" was the meaning of the concept developed in the two beginning statements. This word "cordial" was then used in a question. The three statements preceding the question might be called drill; they could be called training. However, transfer of training in this type of learning takes place in small amounts; it is sometimes negative and again it may be positive. (Woodrow 12). The children were not taught nor trained in generalizing analyzing or abstraction. In such types of learning considerable amounts of transfer take place. (Barlow 1). Thus it is believed that the use of three statements prior to the question was not a learning situation, but each test as presented was merely a verbal problem for the child.

Method of Procedure

The subjects were 15 children ranging from 33.61 months to 56.63 months of age. They attended the nursery school on the campus of George Peabody College for Teachers.

The Play-A-Game tests were given to the children from ten to twelve in the morning. Usually the testing lasted an hour and thirty minutes for four children; time varied for each according to the perseverance of the children taking the test series.

The nursery school teacher introduced the child to the tester, then said: "Would you like to go and play with Miss W. and her toys?" The child took the "Pound a Ball" toy with eagerness. There was no exception. All the children went willingly to play with the toys. In fact it became necessary to send for the children since the tester's appearance in the classroom caused the children to rush forward, clamoring, "Let me go, today".

The Tell-Me test series was more easily administered; not so popular; the time was short; and it varied in length with the willingness of the child to try. One child was given this series in the afternoon, for the tester's convenience. Each child was told that he could play with the toys when he had finished with the stories (verbal tests).

When the Tell-Me tests were given, the children came, listened and answered the questions asked them, but the tester was frequently interrupted in the midst of the story by the child asking, "Where are the toys," or "When can I play with the toys?" All the children sat through the tests except one, who heard the nursery class

TABLE I

SUMMARY OF CORRELATIONS OF CERTAIN FACTORS AS REVEALED IN THE EXPLORATORY STUDY, PERSEVERANCE OF CHILDREN

	Play-A-Game Test Series				Tell-Me Test Series				Test v:s Test		Two Series		Boys		Girls	
			Test				Test									
	Fifteen Children	Twelve Children	Fifteen Children	Twelve Children	Fifteen Children	Twelve Children	Fifteen Children	Twelve Children	Fifteen Children	Twelve Children	Fifteen Children	Twelve Children	Play-A-Game Series	Tell-Me Series	Play-A-Game Series	Tell-Me Series
Tries and Scores	.552	.719	.502	.785	.747	.863	.186	.918					.861	.880	.282	.795
Tries and Ages	-.054	-.001			.152	.011							.557	.358	.285	-.236
Scores and Ages	.392	.336			.120	.237							.858	.310	-.35	.11
Tries and Tries									-.777	.774	-.341	.894	.893		.917	
Scores and Scores									-.465	.813	-.183	.074	.838		.292	

The reliability of coefficients of correlation against the Null Hypothesis by linear interpolation are significant with 15 samples at the 5 percent level when r is .514 and at the one percent level when r is 641; with 12 samples at the 5 percent level when r is .576 and at the one percent level when r is .708. (Garrett 4)

start for a walk on the campus. She said "I have to go now, 'cause they are going for a walk".

The rank order coefficient of correlation was used in calculating the relationship of the factors measured by the tests, since this statistical means is considered efficient for exploratory studies. (Garrett 3:344). However, the number of cases was small, so the rank order coefficients of correlations resulting were tested against the Null Hypothesis (Garrett 4, 298-301).

Results of the Tests

On the Play-A-Game tests the scores ranged from fourteen to thirty-four with a median of twenty-six. The range of total tries was from 179 to 671 with a median of 423 tries. The rank order coefficient of correlation of total scores and total tries was .552. On the Tell-Me test series the range of tries was from zero to eighty-four with a median of thirty-two tries; scores ranged from zero to twenty-six with a median of eight. The correlation between the tries and scores was .765. The corelations of the two test series were: Between tries -.341 and between scores .182. The correlations of age with tries and scores on each of the test series were: On Play-A-Game test series ages and tries -.054 and ages and scores .392: on Tell-Me test series between ages and tries .152 and ages and scores .120.

Since three children from their actions and verbal statements seemed to indicate that they had emotional reactions which blocked their personalities when confronted with a problem, these children's records were removed from the assembled data.

The statistical calculations for the remaining twelve children show: On the Play-A-Game test series the tries ranged from 179 to 671 with a median of 423 tries; the scores ranged from 14 to 34 with a median of 25.5. The correlations between tries and scores was .719, between age and tries was -.001, and between age and scores was .386. The data on the Tell-Me test series show: Range of tries was from zero to 84 with a median of 32; the scores ranged from zero to 26 with a median of eight; the correlations were between tries and scores .853; between age and tries .011 and between age and scores .237. Correlations between the two test series were between tries .894 and between scores .074.

The data show that ages of girls ranged from 42.56 to 48.70 months with a median age of 46.35 months; that age of the boys ranged from 33.606 to 54.7 months with a median age of 44.53 months. When the ages were averaged, the boys were eleven days older than the girls. Three boys were over fifty months old while none of the girls were this old, yet one boy was 33.606 months old.

The girls made a high count in range of tries on the Play-A-Game test Series, 336 to 671, and a median of 468 tries, with a score range from 19 to 34 and a median of 26 score. The correlations were: Between tries and scores .282, between tries and age .285, between scores and age -.35. On the Tell-Me test series the girls made a range of 13 to 84 tries with a median of 34 while the range of scores was from 2 to 26 with a score median of 9. The correlations on this series were: Between tries and scores .796, between tries and age -.36, between scores and age .11. For these girls the correlations between the two series of tests were: Tries .917 and scores .292.

The boys record on the Play-A-Game test series show: Tries ranged from 179 to 499 with a median of 350 tries; scores ranged from 14 to 33 with a median of 22; correlations were, between tries and scores .861, between scores and age, 858, between tries and age .557. On the Tell-Me test series the tries ranged from zero to 63 with a median of 24 tries; scores ranged from zero to 18 with a median of six; for these boys the correlations between tries and scores was .880, correlations between age and tries was .358 and correlation between scores and age was .310. For these boys the correlations of the two series of tests were: Between tries .893, between scores .838.

Data from fifteen children were studied as to each of the seven tests in each of the two series as set up in the order of difficulty. Correlations between the tries was -.777, between scores -.465. However, when the data of three children who had shown emotional confusion during the testing situation were removed and calculations made for the remaining twelve children, the rank order coefficient of correlations were: Tries and tries .774 and scores and scores .813.

Findings

The purpose of the study was to determine the effect of a "child purpose" upon the perseverance of nursery school children in solving objective and abstract problems.

The number of subjects, twelve was small and the results offer only indications to the findings.

Use of psychological tests with a similar study might be indicative of another factor influencing perseverance of young children, who have "a purpose" for solving problems.

Often during the testing in the Tell-Me test series, the child persevered in giving the correct answer to the previous test and only by re-stating the problem did he awaken to the new problem presented. This may indicate perseverance, yet it could be the ability to change from one thought process to another, or the time reaction of the same.

The findings of this simple exploratory study

are summarized as follows:

1. These 15 children did persevere in situations where there was a "child purpose" or play materials which were interesting to them.

The Play-A-Game tests utilized attractive children's toys, which were played with willingly. The children had to be told: "It's time to go now".

For fifteen children the scores ranged from 14 to 34 with a median of 26. Children tried; the range was from 179 to 671 with a median of 423. The rank order coefficient of correlation of scores and tries was .552, which is significant at the 5 percent level; and indicates that relationship exists between scores and tries.

2. Children persevere with academic or language problems, if there is a "child purpose" motivating the accomplishments.

On the Tell-Me test series, only 12 children answered questions and cooperated completely; some wanted to play with the toys they had previously enjoyed. On these tests the scores for the 15 children ranged from zero to twenty-six with median of eight. A perfect score would have been thirty-six. Often a child gave more than one right answer. The answers often came in rapid succession and were counted as tries and scores.

Trying was the number of answers made to the question. Tries ranged from zero to eighty-four with a median of thirty-two The rank order coefficient of correlation between scores and tries was .765 which is significant; showing a relationship between tries and scores with these problems and that the r of .76 is at the .01 level of confidence.

3. Emotional confusion blocks children in problem situations.

Since the tries and the scores were totaled for the tries made by these children on the tests, these correlations are spurious in this exploratory study.

Three children in the two test series became confused and by their overt behavior showed emotional reactions even though they tried and scored in the tests. Examples: Children looked at the tester with unmistakable anger; said: "I don't know," or "I can't do that". This occurred even when they were solving the problem.

When these three children's records were deleted from the study, the data of the remaining 12 children show: On the Play-A-Game test series scores ranged for 14 to 34 with a median of 25.5. Tries ranged from 179 to 671 with a median of 423 tries. Rank order coefficient of correlation between scores and tries was .719. The r has significance beyond the one percent level showing for this small exploratory study relationship exists between trying and scoring.

On the Tell-Me test series these 12 children made scores ranging from zero to 26 with a median of eight. Tries ranged from zero to 84 with a median of 32. Rank order coefficient of correlation of scores and tries was. 853, a significant correlation beyond the one percent level, and showing a relationship between tries and scores.

Also, the deleting of the data of 3 children when attesting the relationship of the tests indicates that confused children cause the testing to go awry. Coefficient of correlation between scores and tries on the two test series changed from -.777 and -.465 to .774 and .813 respectively, which indicate the significance beyond the .01 level of confidence and that have comparable degress of difficulty.

4. In this study with 15 children statistical calculations indicate that age is not a factor involved in perseverance.

Coefficients of correlation were: On the Play-A-Game test series between ages and tries -.054 and ages and scores .392; on the Tell-Me test series ages and tries .152 and ages and scores .120.

For the 12 children, ages correlated with tries and scores as follows: Play-A-Game tests -.001 and .386; Tell-Me tests .011 and .237 respectively. These coefficients are below the level of confidence, and indicating that age of the child was not a factor influencing scores nor tries on these tests.

5. In this study the statistic of treatment of the data indicates that the boys and the girls persevere in practical mechanical problems, yet the girls tried harder and and had less success with objective problems than the boys.

On the Play-A-Game tests the scores of the boys ranged from 14 to 33 with a median of 22, while the scores of the girls ranged from 19 to 34 with a median of 26. The tries of the boys ranged from 179 to 499 with a median of 350; the tries of the girls ranged from 336 to 671 with a median of 468. The coefficients of relationship were: For the girls between tries and scores .282, and for the boys between tries and scores .861. These coefficients are significant for the boys beyond the one percent level and show a relationship between trying and success with these toys. However, the girls tried and did not succeed, the coefficient of .282 shows little, if any, relationship between the scores and tries on these tests. This

may also indicate sex differences in manipulatory skills of these boys and girls.

6. In the situations with verbal problems, stories asking for answers, the girls again tried harder and had less success than the boys. The Tell-Me tests (language problems show: For the girls tries ranged from 13 to 84 with a median of 34, and the scores ranged from two to 26 with a median of nine; for the boys tries ranged from zero to 63 with a median of 24, and the scores ranged from zero to 18 with a median of six. The coefficients of relationship were: For the girls between tries and scores .796; for the boys between tries and scores .880. These coefficients of correlation between tries and scores for these children are beyond the one percent level of confidence. In a hundred trials only once will a lower r appear by accident and indicate relationship between tries and scores.

7. The seven tests in each of the two series, Play-A-Game and Tell Me, were comparable in difficulty.

From the data of the 12 children the coefficients of relationship between the tries on the two series was .774 and between the scores was .813. These r's are beyond the one percent level of confidence, and indicate a comparable relationship of difficulty.

Educational Implications

This simple exploratory study seems to indicate that where there is "a child purpose" children will persevere. Classroom situations at all levels of early childhood education should be created to foster children's purposes in order to develop the habit of perseverance.

In classroom situations every effort should be made to create situations in which any possible emotional confusion is eliminated in order to obtain the best reactions from children.

The age of a child is not likely to be related to the habit of perseverance.

The tests seem to indicate that nursery school girls should have opportunities to use small manual-skill toys, since the girls tried very hard, yet had so little success.

REFERENCES

1. Barlow, M. C., "Transfer of Training in Reasoning," Journal of Educational Psychology, 1827, 28, 122-128.
2. Bott, Helen, "Observation of Play Activities in a Nursery School," Genetic Psychology, Monograph, 1928, 4, 79-87.
3. Garrett, Henry E., Statistics in Psychology and Education, New York, Longmans, Green and Company, 1947, pp. 344-346.
4. Goodenough, Florence L., The Kuhlman-Binet Tests for Children of Pre-School Age, Minneapolis, Minnesota, University of Minnesota Press, 1928.
5. Gutteridge, M. V., The Duration of Attention in Young Children, Melbourne, Melbourne University Press, 1935.
6. Herring, Amanda and Koch, Helen L., "A Study of Some Factors Influencing the Interest Span of Pre-School Children," Journal of Genetic Psychology, 1930, 38, 249-279.
7. Levy, D. M. and Tulchin, S. H., "The Resistance of Infants and Children During Mental Tests," Journal of Experimental Psychology, 1923, 6, 304-322.
8. Recktenwald, L. N., "The Appeal of Information About Occupations," Journal of Experimental Education, 1948, 16, 187-191.
9. Shacter, Helen S., "A Method of Measuring the Sustained Attention of Pre-School Children," Journal of Genetic Psychology, 1933, 42, 339-371.
10. Van Alstyne, Dorothy, Play Behavior and and Choice of Play Materials of Pre-School Children, Chicago, Chicago University Press, 1932.
11. Wolf, T. H., The Effect of Praise and Competition on the Behavior of Kindergarten Children, Minneapolis, University of Minneapolis Press, 1942.
12. Woodrow, H., "The Effect of Type of Training Upon Transfer," Journal of Educational Psychology, 1927, 18, 159-179.

WHAT IS AN I.Q.?

FLORENCE BRUMBAUGH
Hunter College Elementary School
New York City

Introduction

ACHIEVEMENT tests are helpful in diagnosing the academic needs of gifted children but it is difficult to find standardized material which reveals the feelings of such children. Informal opinionnaires, devised by the teachers who seek definite replies to current problems appear to be more valuable than oral queries which may cause the pupils to be influenced by the replies of their classmates. The following account of an attempt to discover the attitudes of pre-adolescent children toward "brightness" may be of interest to others.

All pupils who are admitted to Hunter College Elementary School are required to have a Binet test before entering. These are given by certified psychologists or at educational clinics and in almost all cases prove to be valid measurements of the children's mental ability. The pupils are re-tested when they are in the sixth grade, by the same examiners if possible, and the two scores are compared. It is a rare occurrence when a child whose score was in the 99th percentile at the time of entrance has a lower score when he completes the sixth grade.

The parents receive reports from the examiners that give the percentile rating only, but the actual I.Q.'s are sent to the school. Other intelligence tests are given to the pupils during the intervening years by students at Hunter College and it has been assumed by many persons that the children are aware of their superior intelligence. Discussion by their parents or others, the publicity given to the school might make them conscious of their I.Q.'s, and an attempt was made to discover the degree to which these impressed them by administering a brief questionnaire to seventy-five pupils in three sixth grades a short time before they were to transfer to junior high schools. Following are the directions and questions:

QUESTIONNAIRE

Please do not let anyone know what you write or say except your teacher.

1. Who is the brightest boy in your class? Why?
2. Who is the brightest girl in your class? Why?
3. What does I.Q. mean? Guess if you do not know.
4. Has anyone ever told you what yours is? Write it. Guess if you do not know.
5. What is your greatest problem because you attend Hunter College Elementary School instead of some other school?

Many were bewildered by Question 3, as is shown by the following answers:

Intelligence Quiz (5 replies)
Intelligence Quota (7 replies)
Your speed in tests
How much you can absorb
How high your standard of thinking is
High standards of work and play
Intelligence "equivilment" as against a person's grade
A number that states how smart you are (I think it's 100)
Intellectual capacity
"Mentle standing"
Scholastic standing
Your average in your work
Ability to reason
Brightness
Brilliance
A test you take to get into Hunter and to get out

The remainder of the children wrote Intelligence Quotient, but few spelled both words correctly. The most novel answer was "Ingenuity Quiz" which might describe the child if changed to Ingenuous Quality. Another asked the writer "Do you mean those games in the newspapers and magazines? I never do them because they are so silly." The assumption, based upon these replies, is that I.Q. does not mean much to the eleven-year-olds at this time.

The replies to Question 4 showed such divergence that Tables I and II were prepared for comparison of the known scores with those the pupils guessed. Four pupils stated that they had been told what their I.Q.'s were, but in each case gave incorrect answers. Thirteen made no attempt to guess. Comments were:

I do not know enough about I.Q.'s to gues
I have no idea because I have never seen anyones I.Q.
I don't know how I.Q.'s are marked
My parents never told me exactly, but they say it is very good
I forgot mine but I think it was about 90 (It was 145)

TABLE I

SIXTH GRADE BOYS

Class 1		Class 2		Class 3	
Actual I.Q.'s	Guessed I.Q.'s	Actual I.Q.'s	Guessed I.Q.'s	Actual I.Q.'s	Guessed I.Q.'s
177	150	159	50%	183	?
175	135	153	120	171	99%
170	160	151	140	170	60%
169	99%	148	100	167*	135
164	?	146	130	159	120
164	132	146	98%	156	120
159*	100	139*	145	147	70%
157	140	134	130	142	90%
148	90%			142	135
147	113				
145	146				

*Pupils believed to be the brightest by their classmates.

TABLE II

SIXTH GRADE GIRLS

Class 1		Class 2		Class 3	
Actual I.Q.'s	Guessed I.Q.'s	Actual I.Q.'s	Guessed I.Q.'s	Actual I.Q.'s	Guessed I.Q.'s
175	150	170*	?	178	99%
174	135	167	100	170	114
171	150	165	?	167	97%
170	?	162	99%	165	99%
170	?	159	?	161	105
168	?	153	?	161	60%
167	125	152	97%	157	75%
166	135	149	?	156	132
165	100	149	163	154	110
156*	135	148	90%	153	135
154	140	147	97%	152	150
152	100	146	?	149	150
152	160	145	85%	147	90%
151	130	145	125	139*	125
150	?	143	?	138	69%
145	100				

*Pupils believed to be the brightest by their classmates.

I would say it must be pretty good or I wouldn't have gotten into H. C. E. S.

The wild guesses can be seen by comparing the Binet scores recorded by the psychologists in 1954 with the children's estimates of themselves.

A check was then made to determine whether the pupils regarded the pupils with the highest I. Q. 's as the brightest in their own groups, and in only one case, a girl in Class 2, did these agree. Their reasons for selecting classmates proved to be interesting.

Certain boys were believed to be the brightest because:

Class 1

He reads a lot of good books
Is good at almost anything
Manages to finish what most of us can't
Is good at making friends
Gets good marks
Most of his answers are right
Seems to be more willing to learn than the other boys and doesn't fool around much

Class 2

Brilliant in the science field
Remembers easily
Seems to know what he is doing
Pays attention to directions
Thinks faster than the rest of us
Can snap off facts quickly
Listens

Class 3

Has an answer for everything and is usually right
Thinks out things logically
Has good ideas even if he is careless
He just happens to be smart
Works and if he doesn't know, he tries
Has a good report card
Has a lot of friends

The reasons that some girls were believed to be the brightest were:

Class 1

Gets down to hard tacks and concentrates
Likes tests
Gets along with other children, even the boys
Can figure out examples that I do not know how to start
Good work in class
Gets a lot of "plusses" on her report card
She's nice

Class 2

Can take down notes when the teacher talks
Thinks of what she is going to say before she says it
Studies a lot and knows her stuff
Has no trouble with tests
Is smart and attentive
Just because she is
Is a new girl but just as good as those who have been at H. C. E. S. for years

Class 3

She's always answering questions with a large percentage of them right
She can keep her mind on her work
Remembers everything she hears
Says smart things in class discussion
Listens to directions and works hard too
She has a smart mother so she's smart too

Undoubtedly, personality traits or previous talks by the teachers in regard to qualities that they were trying to stress influenced these explanations. There were scattering votes in all three groups, but in each class one boy and one girl received at least three fourths of the votes, and as they had to list a boy as well as a girl, there was no bias based upon the sex of the pupil as there could have been if the pupils had been asked to name the brightest pupil in the class. One boy wrote, "I must write three names, for they are about equal. Don't think I am bragging, but I think I am one of them." The boy had evaluated himself correctly, but no other pupils listed themselves.

The fifth question was answered by thirty of the pupils with the simple word "None". Twenty mentioned the difficulties due to transportation and twelve that it was hard for them to make friends in their neighborhood. Two mentioned that after they transferred from their local schools their former friends rejected them. Two boys thought that their problem was being dismissed an hour earlier than other schools and thus they were cheated! A boy wrote: "We don't have much homework and in my next school I won't be used to having a lot to do at night." A girl wrote what many others have said at other times: "People expect too much of me just because I go to Hunter. My piano teacher makes me mad because she says I should do better because you have to be bright to go to that school." Another wrote "JEALOUSY. Some of my friends took the test and didn't pass, so now they hate me."

After the papers had been read, the teachers were asked to describe the reaction of the pupils to the questionnaire. One answered, "Resigned," and a child told her, "Another of those things they give us and then we never hear about them again." The second teacher reported that they were relieved when the writer came in to thank the class and told them that they guessed incorrectly and by the use of a yardstick gave them

a slight inkling of what I.Q. meant. The third teacher's comments were that her pupils liked it because it was a different kind of test and they knew they would not be marked for it. One pupil said, "It was not printed so we always know we can say what we want to on those papers.. You have to be careful when there are boxes to mark or you write Yes or No."

Pupils in laboratory schools are subjected to many tests and are seldom anxious about them, but regard them as a nuisance to be endured. It is believed that the responses to the questionnaire were sincere, and that intelligence tests are regarded by these pupils as of no greater importance than achievement tests or personality scales. They realize that every child in the school was given some kind of test before admission, but as the majority entered in nursery school or kindergarten recall it vaguely. Few persons place a premium upon anything possessed by many, and knowing that all of their classmates took the same test and "passed", to them it is not worth discussing. Their admiration is given to those who excel in sports, show special ability in the arts or a particular field, such as science, and above all to those who make friends easily. The greatest number of votes was given to the girl whose I.Q. is lower than any other except one, but she has been popular in every class since she entered at three. The second largest number of votes was given to the girl with an I.Q. of 170 who possesses similar traits. Possibly one could summarize by saying, "If you are a pleasant person, eleven-year-olds think you are bright."

PARTIAL REINFORCEMENT RELATED TO "FREE" RESPONDING IN EXTINCTION WITH PRE-SCHOOL CHILDREN*

NICHOLAS A. FATTU, EDMUND V. MECH, DONAVON AUBLE**
Indiana University

A. The Problem

There exists a lack of systematic data on the human level concerning the effects of partial reinforcement on subsequent response strength during extinction. The experiment reported is concerned with determining the relative efficiency of three schedules of reinforcement in producing superior resistance to extinction when no arbitrary extinction criterion is used during the extinction sessions. In other words, what happens when the organism is "free" to respond as long as he wishes after reinforcement is removed?

B. Design and Procedure

The Ss were assigned to one of three groups randomly with Group A, B, and C having eight Ss each. Each group was given a different condition of reinforcement. Groups A, B, and C were given a series of 12 acquisition trials, with varying conditions of reinforcement, followed by a series of extinction trials. The design of the experiment, similar to that described by Lewis (2), was as follows: "X" signifies reinforcement and "O" signifies no reinforcement.

Group	Reinforcement	Acquisition	Extinction
	%	1 2 3 4 5 6 7 8 9 10 11 12	1 2 3 4 ... n
A	100	X X X X X X X X X X X X	O O O O O
B	50	O X O X O X O X O X O X	O O O O O
C	25	O O O X O O O X O O O X	O O O O O

All Ss were between 3 and 5 years of age They were obtained from the Hoosier Courts Nursery School, at Bloomington, Indiana. Experimental sessions were conducted from 2 p. m. The experimental procedure was carried out as follows. The E entered the nursery school and asked if the children cared to play a new game in which candy could be obtained. No effort was made to persuade those children who did not wish to cooperate, or play. If a child indicated acceptance, he was led to the experimental corner, which was set off from the remainder of the nursery school by a three sided screen 6 ft. in height. Each child was run individually, and when brought into the experimental situation was informed he could either sit or stand before the apparatus.

The E then pointed to a candy pellet in the "food tray," and informed S that more was available from the machine. . Most Ss grasped the idea that more candy would be forthcoming by manipulating the lever or bar on the apparatus. Each S was given a predetermined number of pellets, the magazine having been loaded with 12, 6, or 3 candy pellets depending upon which reinforcement group a particular S was assigned to. The extinction series was initiated immediately after each S had emitted 12 bar presses. Thus all Ss in the 100 percent group received 12 pellets and the Ss in the 25 percent group received 3 pellets. The number of unreinforced responses made by each S to "complete" extinction was taken as the measure of resistance to extinction. That is to say, no time limit was placed upon responding in extinction. Several Ss, during the extinction series thought that the apparatus was broken. They were assured that the apparatus was working properly.

Apparatus. The apparatus consisted of a wooden box, painted green as shown in Figure 1 and was similar to the bar-pressing apparatus described by Siegel and Foshee (3).

In front, at the base, was a plastic food tray. Directly above the food tray was projected a brass rod that served as a bar. A light pressure on the bar closed a switch that energized a release mechanism. A single pellet of candy, 1/4 inch in diameter, was released from a magazine and fed by chute to the tray. Each bar press was recorded by a counter in the control box. The control box was connected to the bar-pressing apparatus by an extension cord permitting the E to move several feet away from S. The control box was constructed so that the E could manipulate the proportion each S would

* Portions of this paper were presented at the 1954 meetings of the American Educational Research Association.

** The authors are indebted to Dr. Wendell W. Wright, Dean, School of Education, Indiana University, for his support.

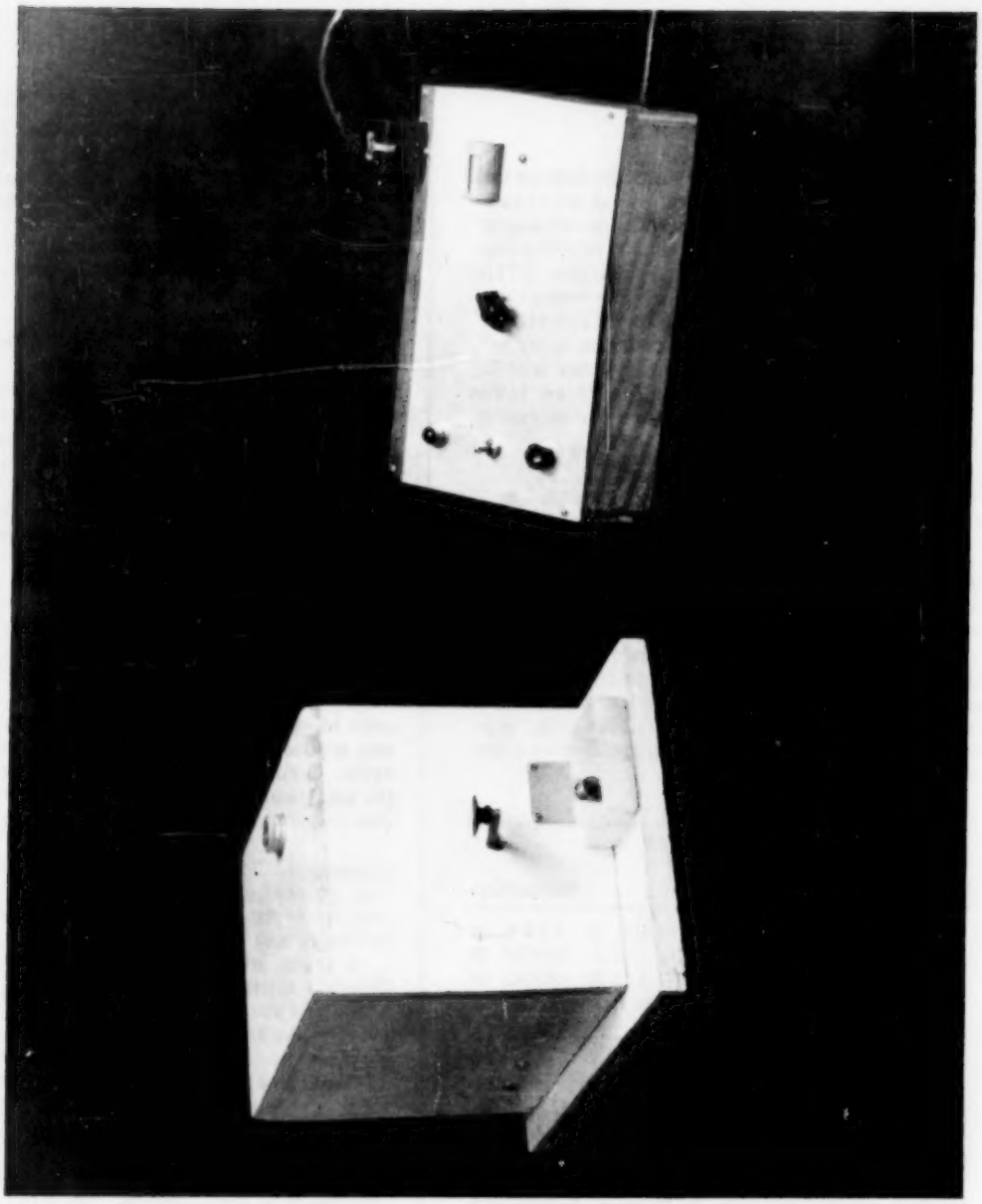
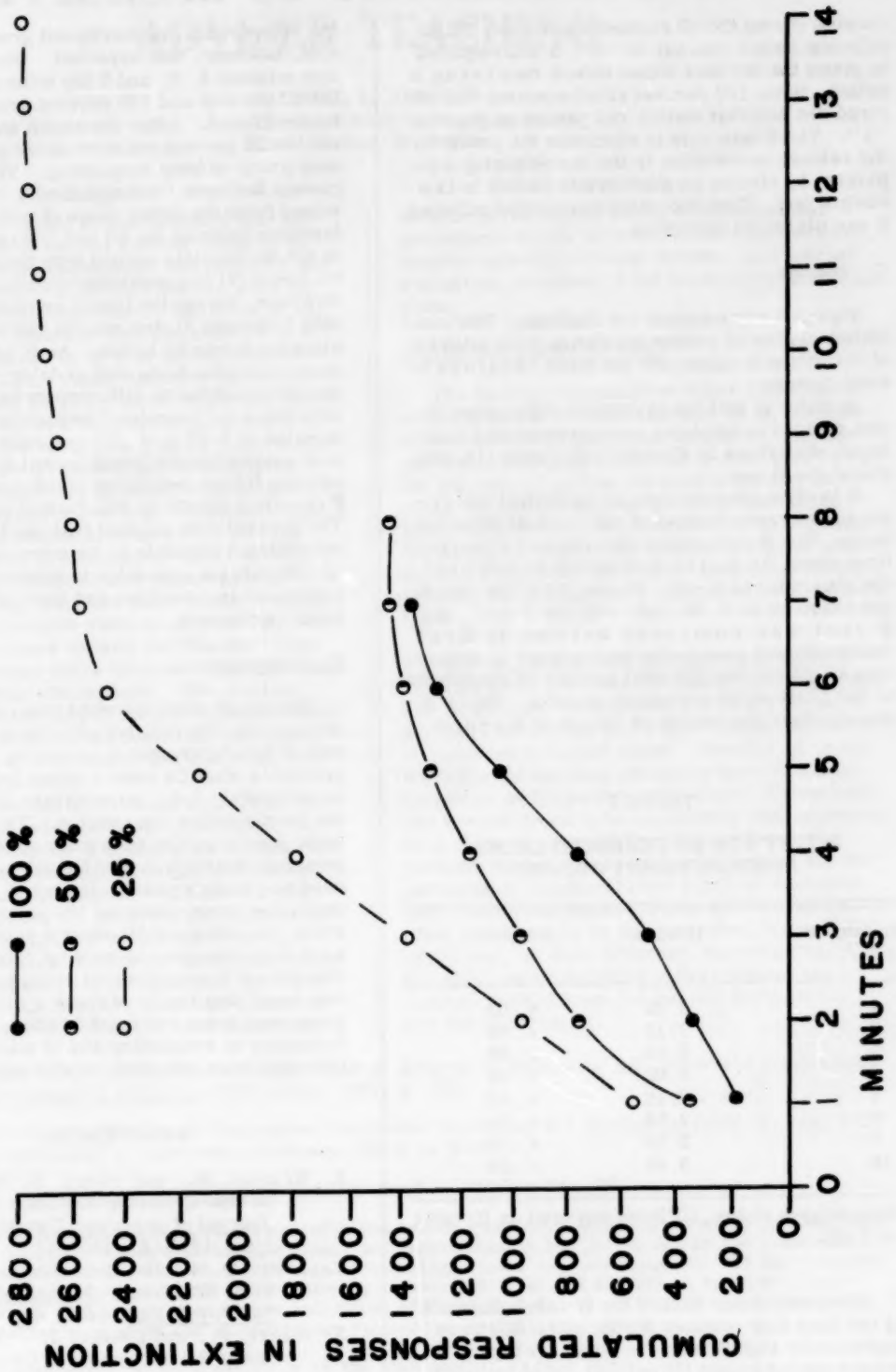


Figure 1

BAR-PRESSING APPARATUS

Figure 2

CUMULATED RESPONSES IN EXTINCTION FOR THE GROUPS RECEIVING 100 PERCENT, 50 PERCENT, AND 25 PERCENT REINFORCEMENT DURING ACQUISITION (GROUP N = 8)



receive during the 12 acquisition trials. If the selector switch was set on "4", S was required to press the bar four times before receiving a pellet. When 100 percent reinforcement was desired the selector switch was placed on position "1". The E was able to eliminate the power to the release mechanism in the bar-pressing apparatus by closing an appropriate switch in the control box. This operation was carried out when S was placed on extinction.

C. Results

Figure 2 summarizes our findings. The cumulated number of responses during each minute of extinction is shown for the three reinforcement groups.

In order to test for statistical differences it was decided to employ a non-parametric technique, described by Kruskal and Wallis (1), and called the H test.

It is often advantageous in statistical analysis to use ranks instead of the original observations. The H test makes only general assumptions about the kind of distribution from which the observations come. Normality of the data is not assumed as is the case with the F test An H test was computed between the three reinforcement groups for each minute in extinction as well as for the total number of responses at the close of the extinction session. Table I summarizes the results of the set of H tests.

TABLE I

H TEST FOR DIFFERENCES AMONG REINFORCEMENT GROUPS*

Minutes in extinction	H-value	P
1	8.82	$< .02$
2	7.12	$< .05$
3	8.20	$< .02$
4	6.05	$< .05$
5	4.10	$< .20$
6	2.88	$< .30$
7-9	3.18	$< .30$
10	3.42	$< .20$

* Significance at the .05 level required an H-value of 5.99.

As evident from Table I the H-values for each of the first four minutes during extinction were statistically significant. In effect, both partial reinforcement groups (50 percent and 25 percent) gave more responses during extinction than did the 100 percent reinforcement group. This result, however, was expected. During extinction minutes 5, 6, and 7 the differences between the 50 percent and 100 percent group ceased to be significant. After the eighth extinction minute the 25 percent reinforcement group was the only group to keep responding. The other two groups had been "extinguished". The data obtained from the latter stage of extinction (no differences between the 50 and 100 percent groups) is not in complete accord with the data obtained by Lewis (2) in a somewhat similar situation. However, during the initial extinction phase (minutes 1 through 4) the results substantiate the conclusions drawn by Lewis. An H test of significance was also computed to determine whether the groups differed with respect to number of minutes taken to "complete" extinction. The obtained H-value of 5.20 ($p < .10$) indicates that the 25 percent reinforcement group tended to respond longer than did the remaining reinforcement groups. Figure 2 points up this finding quite clearly. The present data suggest that the 25 percent reinforcement schedule is superior to either the 50 or 100 percent schedules in producing both more responses in extinction and time taken to "complete" extinction.

D. Summary

The experiment reported was concerned with determining the relative efficiency of three schedules of reinforcement in producing resistance to extinction when Ss were allowed free responding in extinction, i.e., no arbitrary time limit being set for the extinction session. Three groups of eight pre-school Ss each were used, in a bar-pressing situation in which candy pellets were used as reinforcement. During 12 acquisition trials one group received 100 percent reinforcement, one received 50 percent reinforcement, and one received 25 percent reinforcement. The groups then underwent extinction trials. It was found that the 25 percent reinforcement group was more resistant to extinction both in frequency of responding and in minutes to extinction than were the other reinforcement groups.

REFERENCES

1. Kruskal, W., and Wallis, A. "Use of Ranks in One-Criterion Variance Analysis." Journal of American Statistical Association, 1952, 47, 583-621.
2. Lewis, D. "Partial Reinforcement in A Gambling Situation". Journal of Experimental Psychology, 1952, 43, 447-450
3. Siegel, P., and Foshee, J. "The Law of Primary Reinforcement in Children". Journal of Experimental Psychology, 1953, 45, 12-14.

AN APPRAISAL OF TEACHER-COUNSELOR-ADVISER TEACHING

RUSSELL L. JENKINS
Michigan State College
East Lansing, Michigan

FOR THIS study it was assumed that better adjustments to society and our culture, acquisition of skills, knowledge, and attitudes useful in that society and for that culture were desirable educational outcomes. The work of Lippitt and White [1] indicated that the classroom atmosphere had much to do with attitudes and behavior of students. These attitudes and behavior were pointed out to be more frequently friendly to classmates in a democratic atmosphere, more frequently hostile in a laissez-faire atmosphere, and more frequently aggressive in an autocratic atmosphere. A more recent report[2] on changes found to result from autocratic and democratic structured atmospheres (using small samples of five members each, so the authors caution against hasty generalizations from the results) indicates those students in the autocratic atmosphere were more than three times as submissive toward the "leader" than toward the average class member than were those in the democratic atmosphere. This submissiveness indicates lack of initiative, lack of self-insight, lack of individual creativity and expression, and lack of attention to the individual student and his problems.

The experimental study being reported in this paper was designed to seek an appraisal of two instructional methods of teaching communication skills (speaking, writing, reading, and listening). The experimental procedure substituted a democratic atmosphere for the traditional (more authoritarian) atmosphere; substituted "permissiveness"[3] for submissiveness found in the more traditional atmospheres; substituted a "problems-centered"[4] orientation for a problem-centered or subject-centered orientation; and maximized student participation through "discussional"[5] procedures which were substituted for lectures, teacher-question-student-answer, and teacher evaluation practices of the traditional classroom.

Definitions and Procedures

The teacher-counselor-adviser (TCA) method being tested is characterized by an integrated-teacher-role, problems-centered, and experience-providing above and beyond that for the purpose of gaining the knowledge and developing the skills required by the course structure. The procedure is essentially democratic. Groups operate all classroom procedures which would normally be used by an instructor for lecture or teacher-question-student-answer recitations.

Traditional Method and Experimental Method:

Since there are many methods of teaching which are labeled the "traditional method", it seems pertinent to identify more specifically the method referred to in this study. Broadly, it is the method(s) of teaching presently used by the instructors participating in the study. These methods are not found to be essentially different from each other or essentially different from the methods used by other staff members in the Department of Communication Skills at Michigan State College or those used by instructors in other departments at this and other institutions in similar, or even different, departments. The essential characteristics differentiating the traditional method from the method being tested are the following:

1. R. Lippitt and R. White. "Individual Differences as Related to Autocratic and Democratic Group Atmospheres," Psychological Bulletin, XXXVI (July 1939), p. 557.

2. K. Lewin and others. "Studies in Topological and Vector Psychology I," Studies in Child Welfare, XVI (Iowa City: University of Iowa Press, February 1940), p. 307.

3. C. Rogers. Counseling and Psychotherapy (Boston: Houghton Mifflin Co., 1942).

4. The term "problems-centered" as contrasted with the better known term "problem-centered" implies that the class, groups of students within the class, and individuals in the groups and in the class have a myriad of problems facing them in the process of learning. These problems range all the way from emotional and tension problems blocking the learning process and favorable adjustment to other experiences, to those problems related to the individual differences found in ability to master knowledge and learn skills demanded by the standards set for earning credit in a course.

5. The term "discussional" is explained fully in the next section of this paper.

Traditional (LSP)	Experimental (TCA)
1. Essentially autocratic	1. Essentially democratic
2. Major emphasis on student mastering knowledge and skills	2. Major emphasis on problems of individual student
3. Instructor voluntarily gives some formal "instructional" lectures	3. Instructor does not voluntarily give any formal lectures
4. Some time devoted to teacher-question-student-answer recitations	4. No time devoted to teacher-question-student-answer recitations
5. Instructor assumes responsibility and authority for presenting course content material and making individual assignments	5. Class divided into student groups which assume responsibility and authority for presenting course content material and making individual assignments
6. Instructor assumes final authority and full responsibility for all student participation in an evaluation of individual student performance—grading themes and rating speeches; preparing, administering and grading tests and examinations	6. Student groups assume final authority and full responsibility for all student participation in and evaluation of individual student performance during the course, evaluating and giving grades for student themes and speeches. (The instructors were required by the Administration of Michigan State College to assume responsibility for the final grades these students received)
7. Instructor acts as the authority for information given for class acceptance	7. Instructor acts as a resource person for information requested by individual students and/or groups of students
8. Instructor structures class procedures	8. Instructor structures few of the class procedures and participates as a resource person and observer
9. Student participation is not stressed and often little time is devoted to it.	9. Student participation is maximized

The term "less student participation" makes an obvious differentiation of the traditional method from the TCA method being tested. For convenience, this term in the abbreviated form LSP will be used throughout this paper.

Teacher-Counselor-Adviser (TCA):

The role of the teacher is that of a resource person, an observer, stimulator, leader (including the leader of "discussional" procedures, and one who determines the final grade each student will receive in the course). The role of the counselor pertains to, as Erickson[6] defines counselling, "The entire process of helping a person having problems and needs to achieve more desirable goals." The role of adviser pertains to being present at all group meetings, offering any requested help in structuring group activities, and stimulating participation in the group meetings preparatory for group discussions in the class.

Discussional:

As a teacher, the instructor is the leader of discussional procedures. The term "discussional" has been coined deliberately in order that connotative meanings might not get in the way of clear communication. It is important to emphasize that the discussions which were used were not those generally followed in the classroom and have no relation to the commonly used terms such as "discussion", "group discussions", and the like. Two conditions make it mandatory that some term such

6. C. E. Erickson. A Practical Handbook for School Counselors (New York: The Ronald Press, 1949), p. 49.

as "discussional" be chosen to refer to operational procedures which replace the lectures, recitations, and conventional discussions. In the first place, the groups are free to structure their presentations, in most instances, in whatever way they choose. Secondly, the instructor is free to introduce any kind of "democratic" discussion which he feels is appropriate for the situation and the problem at hand. These discussions led by the instructor[7] have been suggested by Phillips and Bradford[8] for use in obtaining greater individual participation and involvement in discussions in problem solving situations. To the writer they seem more democratic than the usual types of discussion in that all students are not just given a chance to participate in the discussions, but in that provision is made such that it is possible for all to participate in the discussion.

Instructors and Testing:

Four instructors participated in the study. Each instructor taught a control class and an experimental class. Sixteen students in each of the eight classes were used as cases for the study. The sixty-four students in the control classes were compared with the sixty-four students in the experimental classes on the basis of scores made on the American Council on Education Psychological Examination and on the basis of the scores made on the Cooperative English Test C 2, and Reading Comprehension (Higher Level) Form Y. These comparisons were made by analysis of variance. No significant differences were found between any two groups or between the control classes and the experimental classes. It could be assumed, then, for the purposes of this study that the variables of ability and reading comprehension were controlled.

The criteria used to measure initial and final achievement in communication skills were those in current use by the Board of Examinations, Michigan State College. These consisted of a three minute speech, a three hundred word paper, and an objective examination. Each student participated in this testing procedure before instruction in the course began, and again participated in this testing procedure after course instruction was finished. These pre- and posttest sources were then used in determining the gain in achievement in the communication skills. The posttest scores were adjusted to the pretest scores and the data then analyzed by analysis of covariance.

Since the TCA instructional procedure was designed to provide training and experience above and beyond that possible with the traditional procedures tested, it was thought desirable to attempt to measure changes other than those in achievement. Student reaction was used to obtain certain subjective data for a qualitative evaluation of the two methods of instruction. The instruments administered to obtain these qualitative data consisted of a sociometric questionnaire (to discover the isolates within the groups at the start of the quarter and the isolates at the end of the quarter), the Student Opinionnaire and the Teacher Evaluation Questionnaire, currently used by Michigan State College to obtain student reaction to course experience and instructor reaction to teaching the course, were administered at the end of the quarter.

Results and Conclusions

Table I summarizes the data bearing on the quantitative data collected during the study. Included in the table are the results of the analysis of covariance to determine significance of the differences noted in the data.

Analysis of covariance does not indicate that either method of instruction is more effective than the other in providing achievement gains in the communication skills. These data do show, however, that under the conditions of this experiment those students in the traditionally taught classes tended to gain more writing skill than did those students in the experimentally taught classes. It seems to indicate, too, that possibly those students in the experimental classes tended to achieve more in the combined skills than did those students in the control classes.

Table II gives a partial summary of the qualitative data. The percent of each of the groups found to be isolate at the end of the first two weeks of the quarter is found in the first column of Table II. Eleven percent of the LSP students were isolate and nineteen percent of the TCA students were found isolate. The fact that the TCA students had worked together in small groups for the two week period might possibly account for some of this difference. The initial period of close contact in the group may have been sufficient to set up stereotype reactions to personalities but not sufficiently long period to make possible better understanding between and among individuals within the groups. The fact that the

7. These procedures consist of the following and variations thereof: Discussion 66, circular response, brain storming. These were developed by Don Phillips, President of Hillsdale College, and were used extensively by him in his leading large group discussions and in his leadership training programs.

8. L. P. Bradford. "Leading the Large Meeting," Adult Education Bulletin (December 1949), pp. 38-50 and 84.

TABLE I

MEANS OF ACHIEVEMENT GAINS

Instructor	Writing		Speaking		Total Skills	
	LSP	TCA	LSP	TCA	LSP	TCA
1	14.50	6.30	3.76	7.67	6.83	6.11
2	8.37	1.83	5.70	0.98	9.05	12.83
3	6.13	-0.08	16.39	16.97	10.05	18.02
4	5.05	0.81	15.36	17.14	10.23	14.92
Ave. Mean Diff. [1]	6.27		0.415		3.93	
F[2]	3.89		0.12		2,35	
Hypothesis[3]	accepted		accepted		accepted	

[1] The highest number of possible points for any student on the writing instrument and also on the speaking instrument was 100. The highest number of possible points for any student on the combined skills instrument was 79.

[2] For the number of degrees of freedom calculated for the statistical procedure used, an F value of 3.95 would be necessary in order for the differences to be significant at the five percent level.

[3] The hypothesis tested is the null hypothesis, that there is no significant difference between method-means when the effects of pre-test scores have been partialled out.

TABLE II

STUDENT-TO-STUDENT REACTIONS AS SHOWN BY A SOCIOMETRIC QUESTIONNAIRE*

	Number of Isolates		Percent Change	Type of Change		
	Pre-test	Post-test		NI to I	I to NI	I to I
LSP	11	20	9	15	5	6
TCA	19	16	- 3	10	12	7

*All data here are given as percent of the total number of students in the classes, LSP data for the control classes, and TCA data for the experimental classes. "I" indicates isolates; "NI" indicates non-isolates. These data were taken from the questionnaire returned by 92 LSP and 90 TCA students.

LSP students were not placed in closely operating groups may have lessened the stereotyping effect of the students in those classes.

The third column indicates the percent of students shown to be isolate at the end of the quarter. Twenty percent of the LSP students were isolates, and sixteen percent of the TCA students were isolate. These changes in percentages of isolates for the two groups indicate nine percent more of the LSP students to be isolate at the end of the quarter, and three percent fewer of the TCA students. These percent changes show that the TCA classes made a reduction in the number of isolates, while the LSP classes made an increase in the number of isolates. This might indicate that the experimental procedures might be influencial in reducing student rejection of peers in the classroom. The close contact of the students operating in the groups may have had some influence in removing the initial stereotyping and permitting students to understand each other better. The counselor role of the instructor may have aided in helping students to adjust to and accept others more objectively. The principle of group therapy may also have had some influence on the changes that came about.

Implications

This study indicates, under the conditions of this experiment, the lecture method (traditionally more autocratic) of teaching produces no better results in communication skill achievement than does the teacher-counselor-adviser method of teaching. A follow-up study might indicate some retained-skill differences for the two methods of teaching.

Each student in the experimental classes spent several hours working with his classmates and for his classmates in an effort to improve communication skills. The group work provided an experience in the laboratory to aid development of group leadership and group membership. These student cooperative and responsibility-assuming activities gave the instructor opportunities to do individual and group counseling within the framework of the course structure.

All of these aids to students, added experiences and training have been shown in this study to be possible without detrimental effects to the "conventional" achievement required by course objectives. This and similar studies may have some value in future teacher training and improved classroom procedures at other levels and for other courses.

PEER GROUP RATINGS OF HIGH SCHOOL PUPILS WITH HIGH AND LOW SCHOOL MARKS

EVAN R. KEISLAR
University of California at Los Angeles

Introduction

THE PROBLEM of this investigation was to discover in what ways high school pupils who get high marks in school differ from equally intelligent pupils who get low marks with respect to the way they are rated, and hence perceived, by their peers of the same sex. Ratings were obtained on twelve traits dealing with social acceptability, approach-avoidance behavior regarding school work, and certain other personal characteristics. While the study makes no attempt to establish causal relationships, such information is likely to give clues as to the nature of social pressures with respect to working for high marks in school, factors which influence teachers in their grading systems, and perhaps personal characteristics related to school success.

The School

The study was conducted among the members of a B11 class in a high school with an enrollment of 1900 in a largely residential section of Los Angeles. The average Otis I.Q., for the class is 104 and the average occupational rating of the fathers is 3.4 on the scale devised by Warner, Meeker, and Eells (4). According to data they present, this is about the level expected of the lower middle class, which would mean that the average pupil in this class comes from a family in which the parental occupational level is higher than about three-fourths of the general population. The population for this study is consequently somewhat selected on both intelligence and socio-economic status.

Measures Obtained

Peer group ratings. —Guess-who ratings on twelve bi-polar traits were obtained for each member of the class from others of the same sex. Out of 209 girls and 227 boys who were technically members of the class, ratings were obtained from 196 girls and 194 boys. For each trait two statements were used, one statement describing behavior on what was arbitrarily defined as the positive side of the trait continuum, the other on the negative; the entire instrument thus consisted of 24 statements. Two forms of the rating blank were constructed, one for each sex. On eleven traits these two forms were identical except for appropriate changes in wording necessary to apply to each sex. Each rater was asked to indicate those class members of his or her own sex whose behavior was described by each statement. Unfavorable ratings were encouraged, first, by asking raters not to sign their names, and, second, by having raters write down assigned code numbers instead of the names of people they wished to mention.

The traits selected are roughly described by paired phrases in Table I, where the first phrase refers to the positive statement and the second to the negative. For many of the statements, the wording was practically identical to that used by Tryon (3). In such cases the same titles have been given to the traits as Tryon used or recommended. For the others a similar form was used. For example, the trait "considerate - inconsiderate" was rated according to the two statements on the boys' form (with appropriate changes in the girls' form being indicated in parentheses) as follows:

Here is someone who is always thoughtful and considerate of other people's feelings and desires:

Here is someone who is not very considerate of other people; he (she) says and does many things which hurt other people's feelings or cause them inconvenience:

Ratings were scored according to the formula given at the top of the next page. This formula yields scores which would be expected if each person were mentioned the same number of times on the total test and thus attempts to control the effect of variations in being widely known. It assumes that the number of total mentions on all the traits is a measure of how widely known in the peer group an individual is since a variety of traits have been used. The number of self mentions, discovered by an invisible fluorescent ink coding device, were so few that they were not excluded in the scoring.

The population from which the high and low marks groups were selected consisted of 128 girls and 126 boys, all of whom had been enrolled at least a full school year and who received 36 or more mentions on the total test. This was done to eliminate those who were newcomers or who were not known well enough to be rated reliably. For this population analyses for the two sexes were carried out separately. The mean scores on the twelve traits for both

$$\text{Trait score} = 50 + \frac{\text{(No. of positive mentions)} - \text{(No. of negative mentions)}}{\text{(Total number of mentions for all twelve traits)}} \times 100$$

sexes (24 distributions) ranged from 48.7 to 53.7, with standard deviations showing a range from 4.8 to 12.6. Since no retest was obtained and since the split-half method was inappropriate, the only evidence of reliability of some of these variables is in terms of amount of agreement between similar ratings. For example, the two traits, liking of school work and putting studies first, showed correlations of .80 for the girls and .78 for the boys in this population. Split-half reliability figures obtained for these trait ratings when scored by the original CEI formula showed fairly typical reliabilities for such an instrument, the median coefficient being .89 for the girls and .86 for the boys.

There is no direct evidence that these ratings really represent how an adolescent is perceived by his peer group. But since with very few exceptions the raters took the test seriously and since every effort was made to keep the ratings confidential, it would appear that such ratings do indicate fairly well the way these adolescents are perceived by those who rated them. The validity of these ratings in terms of their agreement with other independent estimates of such characteristics is not established. However, since the two traits just referred to showed fairly good agreement with school marks, as indicated in Table I, we have some evidence that at least here the ratings agree with an independent criterion. But it is probable that for most of the other traits in the instrument either greater halo effect was operating or the raters had had less opportunity to make judgments.

School Marks. —The average school mark received during each pupil's previous year in high school, the tenth grade, was translated into a number where A = 9, B = 6, C = 3, and D and F = 0, with plus and minus values receiving intermediate figures. The mean grade received was 4.8 with a standard deviation of 2.1 for the 128 girls, and 4.2 with a standard deviation of 2.5 for the 126 boys.

Subjects

From the 128 girls described earlier, two groups of 27 pupils each were formed by selecting pairs of girls who had approximately the same Otis I.Q. but differed by at least 3 points in school marks (one letter grade). In a similar fashion two groups of 35 boys each were formed from the population of 126 boys. For each sex, therefore, there were available a high marks group and an equally intelligent low marks group. The effectiveness of matching may be judged by noting that the girls with high and low marks showed mean I.Q.'s of 105.8 and 104.7, respectively. For boys the comparable figures were 106.5 and 106.9. Standard deviations in I.Q. for these groups were all approximately 12 points. The mean grades for the high and low marks groups were 7.0 and 2.2, respectively, for the girls, and 6.7 and 2.0 for the boys. These two groups for each sex were then compared with respect to the ratings received.

Results

We may first note some general relationships obtained from the populations of 128 girls and 126 boys described earlier. As indicated by Table I, fairly high correlations were found between school marks and each of the three variables, liking of school work, putting studies first, and persistence for both boys and girls. Less pronounced but significant relationships are to be noted between school marks and considerateness (for both sexes), being influential and friendliness (for girls) and inversely for conceit (for boys). While the same general pattern is found between intelligence and the ratings, the relationships are much lower than in the case of school marks. For boys, occupational level of father was positively correlated with liking of school work (.31) and negatively with competency in athletics (-.27). Otherwise none of the correlations of the ratings with this variable was significant for either sex at the .01 level and thus have not been presented.

The results of the comparisons between the two groups matched for intelligence are given in Table II for the girls and Table III for the boys. The value of s_D for each comparison was computed from a distribution of differences between the matched members of each pair on the trait in question. The hypothesis of normality of these trait distributions obtained from the total population of each sex in this study could not be rejected at the .05 level except for the two traits indicated by footnotes. The results of the t test show that the girls with high marks differ significantly from equally intelligent girls with low marks in that the former were rated as liking school work more, more persistent, more likely to place their studies first, more influential and less popular with the opposite sex. With

greater uncertainty we may state that the girls with high marks were also rated as more considerate, more friendly, less talkative, and less conceited. As compared with boys with low marks, equally intelligent boys with high marks were rated as liking school work, more persistent, more likely to give primacy to their studies, and more considerate. With greater uncertainty we may conclude that boys with high marks were also rated as acting older and being less conceited.

The three traits, liking of school work, putting studies first, and persistence, yielded the highest correlations with school marks and showed the clearest differences between the high and low marks groups for each sex. These findings do not necessarily indicate a high relation between motivation or work habits and achievement. This is borne out in part by the fact that the partial correlations of intelligence with each of the traits, liking of school work and putting studies first, when school marks are held constant, are about zero, -.07 and -.05 for the girls and .13 and -.04 for the boys, instead of being substantially negative as one would expect. In other words, bright pupils who received C's were rated about the same on these two traits as dull pupils who received C's. It is likely that raters were influenced by the marks pupils had received in making such judgments or that teachers are influenced by such factors in assigning grades, or perhaps both.

In general, for boys there appeared to be little relationship between school marks and the traits which are related to social acceptability, namely, friendliness, popularity with opposite sex, and being welcomed by the same sex. Not only were no differences found between the high and low marks groups on these traits, but the correlations for the entire population between these traits and school marks were not significantly different from zero. When the possibility of a curvilinear relationship was tested by studying the regression of each one of these three traits on grades, none of the η's was significant. However, the sum of the scores on these three traits yielded an η of .37, which was just significant at the .05 level. ($F = 2.07$, $n_1 = 9$, $n_2 = 116$.) Subsequent t tests showed that those boys who received the average grade of C+ were rated higher (differences significant at the .01 level) on these three traits than those who received grades of D+ and lower or those who received A- and higher. But with such a low level of significance for η, real differences among the array means are not satisfactorily established. We may, therefore, conclude only with considerable uncertainty that boys who get a "gentleman's" grade of average are rated higher on these three traits related to social acceptability than are boys with very low marks or those with very high marks.

Girls with high marks are rated as being significantly more influential than equally intelligent girls with low marks. but on the more personal traits of being welcomed and friendly the differences, while in favor of the high marks group, are not significant at the .01 level. The finding that girls with high marks are rated as significantly less popular with the boys may indicate a stereotype on the part of the raters, the presence of some compensatory device, or something about the characteristics of girls which boys in this population prefer. No evidence of lack of linearity was found for the relationship of any of these traits, either singly or combined, to grades among the entire population of 128 girls.

Discussion

While for boys there appears to be no linear relationship between school marks and traits related to social acceptability, the unreliable finding that boys with average marks tend to be rated higher on a combination of such traits than those with high or low marks is worth further investigation. Of interest also is the fact that girls with high marks are rated as less popular with the opposite sex than are equally intelligent girls with low marks. This is in contrast to the finding that girls with high marks are rated as significantly more influential (at the .01 level) and more friendly (at the .05 level). As rated by their own sex, girls with high marks tend to receive different ratings with respect to their relations to the two sexes.

If we may assume that these peer group ratings are valid estimates of certain traits possessed by the pupils, the results of this study may be interpreted to support the hypothesis that teachers are influenced by personality factors in giving grades. Apart from the close relations of such factors as liking of school work, persistence, and putting studies first, we may note that the group with high marks in each sex was rated as more considerate and less conceited; girls with high marks were rated as less talkative, and boys with high marks were rated as older acting, although many of these differences are significant at only the .05 level. As such, these findings in part agree with teachers' ratings of pupils (2, 5).

On the other hand, it is possible that such differences reveal the characteristics of pupils who possess a high anxiety related to doing well in school. Those pupils who have acquired such motivation may come from homes, as Allison Davis suggests (1), where other traits are also fostered. The finding that pupils with high marks are rated not only higher on liking school work and putting studies first, but higher on

persistence and considerateness, and lower on conceit may suggest a pattern in which the social anxiety of this group is expressed. While according to Allison Davis' thesis, social anxiety is characteristic of children from the middle class, no relationship was found between these ratings and socio-economic status. Such negative findings may have been caused by the restricted range in occupational level of the population in this study.

Summary

For each sex a group of pupils with high marks was compared with an equally intelligent group of pupils with low marks on twelve bi-polar trait ratings obtained from the members of their own eleventh grade class of the same sex by means of the guess-who technique. For both sexes the high marks group was rated higher on liking of school work, putting studies first, and persistence. In addition, the girls with high marks were rated as being less popular with the boys and more influential. For the boys the high marks group was rated as being more considerate. Other differences, significant at only the .05 level, were found. Certain hypotheses were suggested to account for the findings.

REFERENCES

1. Davis, Allison. "Socialization and Adolescent Personality," Forty-Third Yearbook, N.S.S.E., (Chicago: University of Chicago Press, 1944), pp. 198-216.

2. Leipold, L. E. "Who Are Our Good Students?" Journal of Educational Research, XXXVIII (1945), pp. 529-533.

3. Tryon, Caroline McCann. "Evaluation of Adolescent Personality by Adolescents," Social Research Monograph on Child Development, No. 4, Serial No. 23 (1939).

4. Warner, W. Lloyd and others. Social Class in America (Chicago: Science Research Associates, 1949), pp. 274.

5. Wolf, S. Jean. "Historic Background of the Study of Personality as it Relates to Success or Failure in Academic Achievement," Journal of Genetic Psychology, XIX (1938), pp. 417-436.

REGRESSION COEFFICIENTS VS. CORRELATION COEFFICIENTS*

JOHN SCHMID, JR.**
Lackland Air Force Base

Introduction

RESEARCH IN the area of human factors frequently calls for determining relationships among characteristics of individuals. That is, we are often interested in determining if there is a linear relationship between two variables, such as success in school and number of hours spent in study, intelligence and school achievement, flying proficiency and amount of previous flying experience, success in some Air Force technical training school and a screening test for admission to that school, ...

The usual approach in determining the existence or extent of this relationship has been to calculate a coefficient of correlation between the two traits under investigation. Casual examination of the literature seems to indicate that alternate and, perhaps more appropriate indices are seldom used. Although, in many cases, the product-moment correlation coefficient may be a satisfactory index of linear relationship, there are occasions when it might be more informative to use a regression coefficient to describe the relationship. The widespread use of product-moment formula and consequent disregard of regression probably underlies Fisher's statement:[1] "The idea of regression used usually to be introduced in connection with the theory of correlation but it is in reality a more general and simpler idea. Moreover, the regression coefficients are of interest and scientific importance in many classes of data where the correlation coefficient, if used at all, is an artificial concept of no real utility".

It is my purpose to discuss a few characteristics of regression coefficients which, for certain purposes, make them more useful than the corresponding correlation coefficients.

Direction of the Relationship

In examining the linear relationship between two variables, frequently we are not interested in only a numerical value representing the strength of the relationship but, rather, how changes in one variable are accompanied by changes in the other variable. For example, the Air Force may be interested in determining if a certain type of paper-pencil test may validly be used for selecting men for flying training. Here, we are attempting to predict flying success from test scores and the reverse prediction of test scores from flying success would be quite meaningless. Prediction of one characteristic from knowledge of another characteristic injects a temporal factor into the relationship such that the the relationship may be considered as a uni-directional rather than a two-way construct. That is, we attempt to find out whether a certain degree of proficiency on the test, independent variable, is followed by the consequence of a certain degree of flying proficiency, dependent variable; and we are not at all interested in the reverse sense of this relationship. When a unidirectional relationship is inherent in the problem, as here, a regression coefficient is an appropriate index to use, whereas a correlation coefficient is more appropriate for those problems implying a two-way relationship between the traits. It is this distinction in function between regression and correlation coefficients which identifies a consideration too often overlooked. It is incumbent upon a research worker to give due attention to the types of variables which he is investigating, the questions which he wishes to answer, and the sense of the relationship with which he is concerned. These are considerations which call for attention and forethought rather than indiscriminate application of the product-moment formula. Whenever the research worker is attempting to estimate the value of one variable from knowledge of the corresponding value of the other variable, then a regression of the dependent variable upon the independent variable is the statistic that furnishes the information desired. On the other hand, when he is interested in determining a two-way relationship between the variables, the correlation coefficient is the appropriate statistic to be used. Use of regression, then, has the natural advantage of compelling the research worker to think more carefully about the variables he is working with; that is, whether he is predicting or estimating a trait and what trait he will use to do this predicting or estimating.

* The substance of this article was presented at a Proficiency Measurement Conference held at Headquarters, Crew Training Air Force, Training Analysis and Development, Randolph Air Force Base on August 12, 1953.

** Supervisor of Statistical Processing Branch, Research Services Division, 6560th Research and Development Group.

1. Fisher, R. A. Statistical Methods for Research Workers. London: Oliver and Boyd, 1950: pp. 128-9.

TABLE I

RESTRICTION OF RANGE INDICES OF RELATIONSHIP BETWEEN AGCT SCORES AND A AND E JET MECHANICS SCHOOL GRADES

Degrees of Restriction	Correlation Coefficient	Regression Coefficient
0%	.57 (1)	.022 (2)
11%	.55	.024
20%	.52	.025
36%	.44*	.022
55%	.43*	.025
72%	.34*	.020
(1) 5% fiducial limits	.50 to .64	
	− +	
(2) 5% fiducial limits	.019 to .025	

*Significantly different from the population correlation of .57.

Preservation of the Metric

A second advantage in using regression is that it preserves the units of the variables being measured. A regression coefficient tells us how many units change occurs in the dependent variable with a unit of change in the independent variable. In some problems, this knowledge might be relatively trivial, but in other problems it may be quite important. For example, it might be quite valuable to know that men who earned raw scores of 83 or higher on a motor skills test commit 15 fewer errors, on the average, in a particular recovery maneuver in flying training than men who score less than 83. These units may have concrete meanings and useful implications for the people attempting to predict flying success from a previously administered motor skills test. If a correlation coefficient had been used to indicate this relationship, the units for both variables would have disappeared, leaving only a statement of the strength of the relationship. Perhaps, the value of using an index of relationship which preserves the original units of each variable is more clearly seen if we consider the hypothetical problem of the relationship of pounds of a potash type of fertilizer to the yield of wheat. Generally, telling a farmer that the relationship between wheat yield and the fertilizer is, say .56 is equivalent to telling him nothing. However, if the information is expressed as a regression coefficient, that is, if he is told that wheat increases about 2.7 tons per acre with each 100 pounds application of fertilizer, then he is presented with useful and meaningful facts.

Effects of Selection

A third advantage of regression coefficients over correlation coefficients, for certain problems, lies in the effect of selection upon correlation. A regression coefficient does not depend upon the frequency distribution of the independent variable. That is, values of the independent variable may be chosen in any arbitrary manner and the regression slope will be unaffected, whereas arbitrary selection may severely deflate the numerical value of the coefficient of correlation. In the selection and classification of Air Force personnel, we are confronted with the problem of determining relationship between technical school grades and certain screening tests. Selection of the top 60 or 70 percent of men for a certain technical school course by means of a screening test results in a seriously attenuated correlation coefficient between school grades and the screening test scores. The following table shows how various degrees of selection affects the coefficients of correlation, but leaves the corresponding regression coefficients relatively unchanged.

Considering the coefficient of correlation and regression coefficient of the 0 percent restricted group, i.e. the unrestricted group, as population values, we may set up fiducial limits which represent the extent to which corresponding statistic values may be expected to fall 95 times out of 100 cases upon random sampling. We see that as the percentage of restriction increases, the coefficients of correlation decrease from the population value of .57 to .34, whereas the regression coefficients fluctuate between .020 and .025. The fiducial limits for the coefficient of correlation given at the bottom of the table show that when restriction has occurred to the extent of at least 36 percent the coefficients of correlation resulting tend to deviate further than sampling would permit. That is, the restriction for 36, 55 and 72 percents results in coefficients of correlation which ordinarily would be considered as different from their corresponding population value of .57. It is true that we do have correction formulas which will estimate the population value if we know how much restriction has occurred. However, all of this is obviated by use of the regression coefficients. We see that even in the case where restriction has been as much as 72 percent the regression coefficient does not differ from its population value more than what may be expected on the basis of chance.

Summary

In summary, then, we can say this. Frequently it is more informative to use a regression coefficient rather than a correlation coefficient to describe a linear relationship. When a research worker can, on logical grounds, say that he is concerned with the uni-directional relationship between two variables, a regression coefficient may be appropriately used. Secondly, regression coefficients preserve the original units of the related variables. Thirdly, whereas coefficients of correlation are seriously affected by various types of selection on the independent variable, this is not true of the regression coefficient.

Specifications for Manuscripts

. . . for the . . .

JOURNAL OF EDUCATIONAL RESEARCH

. . . and the . . .

JOURNAL OF EXPERIMENTAL EDUCATION

1. All manuscripts must be typewritten, double spaced, and on one side of the sheet only. Mimeographed and ditto sheets are acceptable only when very clearly printed.

2. All unusual symbols or formulae must be very clearly typed or hand printed in black ink. To avoid costly printers' composition charges it may be necessary for us to make cuts of difficult matter, or to print your material by the photo-offset lithography method. The latter means photographing your actual copy. It is expensive to have material redrawn by our own artists, and retracing or duplicating increases the hazards of error. See that your copy is correct and complete as you wish to have it reproduced. The men who work on your manuscripts are not trained to understand the working symbols and language of your technical field.

3. The same restrictions and requirements as in Paragraph 2 apply to all drawings, graphs or other illustrated materials,—they must be neatly done, in black ink, on bond paper or tracing cloth suitable for reproduction. Remember our magazines are printed in black ink only. Color graphs should be changed by the author to provide different kinds of shading for the different areas. For example: diagonal lines for red, vertical lines for blue, etc. Provide a key.

4. All tables, graphs, etc., on sheets by themselves must be properly labeled and identified in relation to the written copy of the manuscript.

5. Footnotes must be complete as to author, title, place of publication, publisher, date and pages. They must be numbered consecutively throughout the article.

6. Bibliographical notes must be complete and arranged alphabetically.

The cooperation of all prospective authors in following these rules is earnestly required. It is difficult to produce technical journals accurately, neatly, and on time under the best conditions. Promptness in printing, economy, and accuracy will be promoted by carefully prepared manuscripts.

A New

Educational Statistics Primer

Written by students for students

- Statistics self-taught for the non-mathematical student
- Who desires a more intelligent reading of technical papers in education, and
- Who desires to make easy statistical calculations with speed and accuracy

This pocket edition of statistics explains and illustrates calculation procedures for:

Measures of central position

Measures of variability

Zero order correlation and simple regression

Errors of measurement and sampling

Multiple correlation and regressions

Ready now . . . . . Price $1.25

DEMBAR PUBLICATIONS, Inc.

Madison 3, Wisconsin

100 MILLIMETERS

INSTRUCTIONS Resolution is expressed in terms of the lines per millimeter recorded by a particular film under specified conditions. Numerals in chart indicate the number of lines per millimeter in adjacent "T-shaped" groupings.

In microfilming, it is necessary to determine the reduction ratio and multiply the number of lines in the chart by this value to find the number of lines recorded by the film. As an aid in determining the reduction ratio, the line above is 100 millimeters in length. Measuring this line in the film image and dividing the length into 100 gives the reduction ratio. Example: the line is 20 mm. long in the film image, and $100/20 = 5$.

Examine "T-shaped" line groupings in the film with microscope, and note the number adjacent to finest lines recorded sharply and distinctly. Multiply this number by the reduction factor to obtain resolving power in lines per millimeter. Example: 7.9 group of lines is clearly recorded while lines in the 10.0 group are not distinctly separated. Reduction ratio is 5, and $7.9 \times 5 = 39.5$ lines per millimeter recorded satisfactorily. $10.0 \times 5 = 50$ lines per millimeter which are not recorded satisfactorily. Under the particular conditions, maximum resolution is between 39.5 and 50 lines per millimeter.

Resolution, as measured on the film, is a test of the entire photographic system, including lens, exposure, processing, and other factors. These rarely utilize maximum resolution of the film. Vibrations during exposure, lack of critical focus, and exposures yielding very dense negatives are to be avoided.